The Magic of ACT-R and the Importance of Noise in Learning
Jim Treyens - 1/6/2020

For Andrea's seminar last quarter I put together a simple model of the Iowa Gambling Task (IGT) (Bechara et al., 1994). The scorecard below is from a typical control subject:

RESPONSE OPTION	1	2	3	4	5	6	7	8	9	10
A +100	9	10	11 -150	29	30 -300	47	48 -200	57	58 -250	93 -350
B +100	1 0	2 0	3	18 0	19	20 0	21	22	36 -1250	49
C +50	6	7	8 -50	23	24 -50	25	26 -50	27	28 -50	33 -50
D +50	4 0	5	12 0	13	14	15 0	16 0	17	31	32 -250

RESPONSE OPTION	1	2	3	4	5	6	7	8	9	10
A +100		-350		-250	-200		-300	-150		
B +100	50 0	51 0	52	53 -1250		0	0			
C +50	34	35 -25	63 -75	64	65	66	67 -25	68 -75	69	70 -50
D +50	37 0	38 0	39	40	41 0	42	43	44	45 0	46 -250

RESPONSE OPTION	1	2	3	4	5	6	7	8	9	10
A +100		-300		-350		-200	-250	-150		
B +100	-1250		0	0		0		0		
C +50	71	72	73	74 -50	75 -25	89 -50	90	91	92 -75	94 -50
D +50	54	55 0	56	59	60 0	61 0	62 0	76	77 -250	78

RESPONSE OPTION	1	2	3	4	5	6	7	8	9	10
A +100	-350	-200	-250				-150	-300		
B +100	0	-1250		0		0		0		
C +50	95	96	97	98 -25	99 -25	100	-75		-50	-75
D +50	79	80 0	81	82 0	83 -250	84 0	85	86	87 0	88

Notice that typical control subjects learn relatively quickly that the payoff (rewards – penalties) is better for decks C and D.

I built a simple reinforcement model in ACT-R for doing the IGT. I set :alpha (learning rate) at 0.2 (which is the ACT-R default) and :egs (expected gain s which specifies the s parameter for noise added to utility values) to 5. This produced a scorecard that looks very different from the representative control scorecard: decks A and C (with lower penalties) are chosen very often and decks B and D (with higher penalties) are chosen much less often, e.g., see below:

Trial 1

Pick	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Deck A + 100	2	4	5	6	7	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79
			-150		-300		-200		-250	-350		-350		-250	-200		-300	-150		
Deck B + 100	44	45	46	47	48	49	50	51	52	100										
									-1250					-1250						
Deck C + 50	1	3	8	10	11	12	13	14	15	16	19	20	21	23	24	27	29	30	34	35
			-50		-50		-50		-50	-50		-25	-75				-25	-75		-50
Deck D + 50	9	17	18	22	25	26	28	31	32	33										
										-250										-250

Pick	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Deck A + 100	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99
		-300		-350		-200	-250	-150			-350	-200	-250				-150	-300		
Deck B + 100																				
	-1250											-1250								
Deck C + 50	36	37	38	39	40	41	42	43	53	54	55	56	57	58	59	60	61	62	63	64
				-50	-25	-50			-75	-50				-25	-25		-75		-50	-75
Deck D + 50																				
									-250						-250					

Now here's the magic: I changed the :egs (noise) parameter to 23 and now the model very closely matches the results of the typical control subject (no brain damage):

Trial 2

Pick	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Deck A + 100	23	26	29	31	34	43	45	58	61	71	82									
			-150		-300		-200		-250	-350		-350		-250	-200		-300	-150		
Deck B + 100	2	3	4	5	6	7	8	9	10											
									-1250					-1250						
Deck C + 50	13	16	18	25	27	30	32	35	36	37	38	39	41	46	49	53	55	57	62	64
			-50		-50		-50		-50	-50		-25	-75				-25	-75		-50
Deck D + 50	1	11	12	14	15	17	19	20	21	22	24	28	33	40	42	44	47	48	50	51
										-250										-250

Pick	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40
Deck A + 100																				
		-300		-350		-200	-250	-150			-350	-200	-250				-150	-300		
Deck B + 100																				
	-1250											-1250								
Deck C + 50	66	67	69	70	73	74	75	77	78	79	80	81	83	87	88	89	90	98	99	100
				-50	-25	-50			-75	-50				-25	-25		-75		-50	-75
Deck D + 50	52	54	56	59	60	63	65	68	72	76	84	85	86	91	92	93	94	95	96	97
									-250						-250					

I find it amazing that a simple ACT-R reinforcement model can so closely model the behavior of normal control participants without using declarative memory at all.

It is also clear that noise is essential in facilitating learning. When noise is too low, the model locks on to Decks C and A which have smaller penalties and quickly abandons Decks B and D, which have infrequent but larger penalties. In two of the five trials I ran, Deck D is not even sampled at all. Below is a summary of the five trials with low noise. More detail on the five low-noise trials is in IGT_Summary_Reinf_alpha0_2_egs_5.pdf

IGT reinforcement learning model with alpha = 0.2 and egs = 5

	Picks per Deck			
	Deck A	***Deck B***	***Deck C***	***Deck D***
Trial 1	40	10	40	10
Trial 2	40	20	40	0
Trial 3	40	20	40	0
Trial 4	30	9	40	21
Trial 5	40	9	20	11
Mean	38	13.6	36	8.4

In the trials with higher noise, the model quickly learns that Decks C and D have the higher return. The results for the five trials below with higher noise are remarkably consistent and are similar to the typical control subject result presented in Bechara et al., 1994. More detail on the five higher-noise trials is in IGT_Summary_Reinf_alpha0_2_egs_23.pdf

IGT reinforcement learning model with alpha = 0.2 and egs = 23

	Picks per Deck			
	Deck A	***Deck B***	***Deck C***	***Deck D***
Trial 1	11	9	40	40
Trial 2	11	9	40	40
Trial 3	12	9	40	39
Trial 4	12	9	39	40
Trial 5	11	9	40	40
Mean	11.4	9	39.8	39.8

Now my task is to build a memory model for the IGT. It will be interesting to see whether a memory model will be as good at modeling typical control subject results. However, I suspect it will be necessary to use a memory model in order to understand brain-damaged patients' results. The fact that the simple reinforcement model is so good modeling normal subjects may make it more difficult to reject that model entirely.

The reinforcement model for IGT used for this note and related files can be found at https://github.com/jtreyens/IGT

Reference

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1-3), 7-15.