Computational Models of Decision Making:

The Effect of Computational Complexity on Neural Reaction Time

---

Peter Grabowski

Princeton University

Key Points

1. Model Free Decisions: Habitual or instinctual, no reasoning or planning. Computationally Simple
2. Model Based Decisions: Made based on mental model. Computationally complex; think Djikstra's algorithm or depth first search
3. Daw's 2011 study allows us to quantify how "model based" or "model free" someone by creating computational models based on reinforcement learning
4. Current understanding predicts processing time increases with computational complexity → model based decisions should take longer than model free ones
5. Assess correlation between reaction times and *w,* the computational model's representation of how "model based" a subject was

Abstract

Recent research conducted by Nathaniel Daw (Daw et al, 2011) has suggested that two key processes underlie human decision making. The first, referred to as "model free" decision making, encompasses choices that are instinctual or habitual, while the second, known as "model based" involves choices that are planned or reasoned. The current study combines Daw's findings with theories of computational complexity to make predictions about the correlation between the prevalence of model based decision processes in a subject and their average reaction time. Current theory predicts that as computational tasks become more complex, the time required to complete them increases. An analysis of the logic behind model free and model based decision making reveals that model based decisions are much more computationally complex. As such, if the laws governing computational complexity and processing time can be applied to neural circuitry in the same way they are applied to silicon circuitry, subjects who are more model based in their decision making should have slower reaction times. Our study replicates and extends Daw's finding in an attempt to determine the relationship between the relative strength of model free and model based decision making processes displayed by a subject and how quickly they responded

A State-Action Diagram

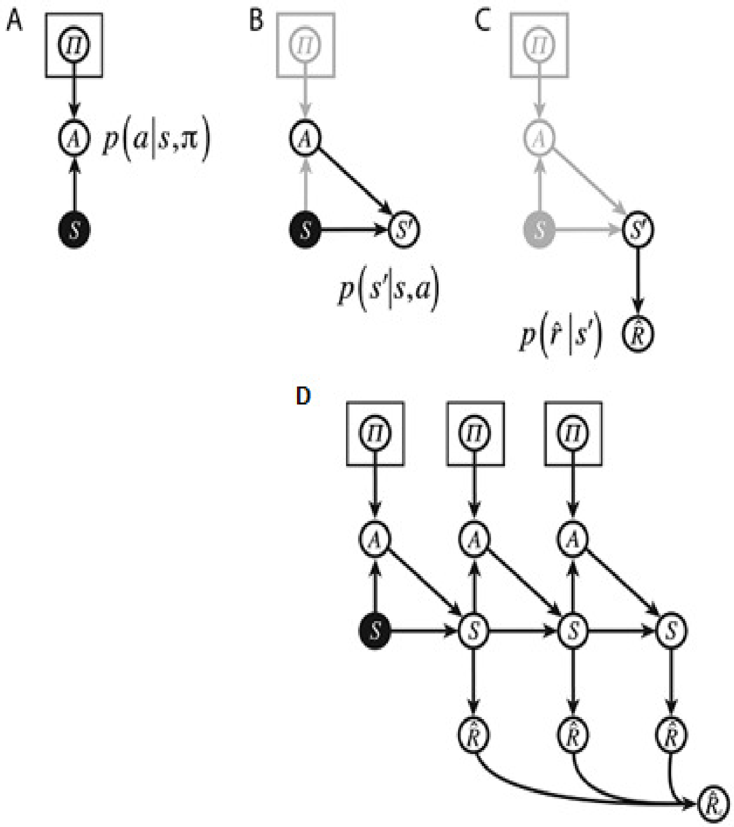


Figure A: Given a state *s* and a policy *π* with which to evaluate it, the agent can select an action *a*.

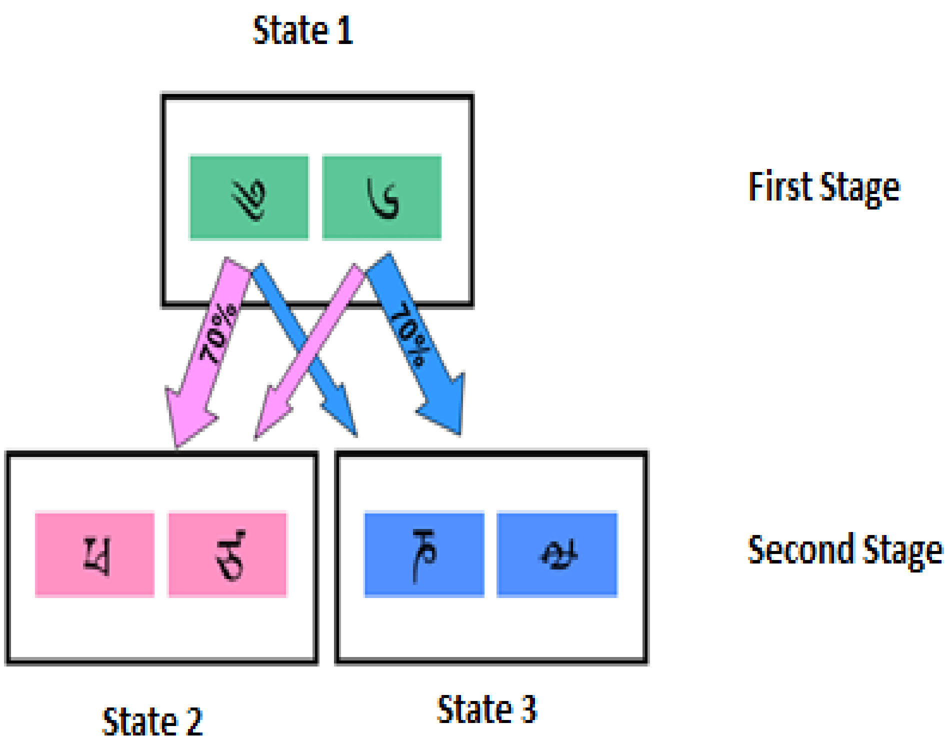
Figure B: Given a state *s* and action *a*, the agent finds itself in a new state *s'*.

Figure C: After finding itself in a new state *s'*, the agent receives a reward $\hat{r}$.

Figure D: This process can be iterated over, in each case feeding the new state *s'* generated by the selection of action *a* under policy π to come to a new state *s*'', resulting in a new reward $\hat{r}$. These rewards are summed over all the iterations to produce reward *R*. For all processes described in Figures A-D, the underlying selection mechanics can be either stochastic or deterministic

Adapted from Solway and Botvinick, 2012

Daw's Two Stage Decision Task



Each trial consisted of two stages. In the first stage, subjects were presented with two choices, labeled by semantically irrelevant Tibetan characters. Choice of the left character (in this depiction) led to State 2 on seventy percent on trials (the "common" transition) and to State 3 on thirty percent of trials (the "rare" transition). The opposite held true for an initial choice of the right character. In the second stage, subjects entered into one of two States (State 2 or State 3), depending probabilistically on their first stage choice. The probability of a reward after a choice in the second stage varied, and was updated after each trial by a random Gaussian walk.

Adapted from Daw et al., 2011

Expected and Actual Stay Probabilities

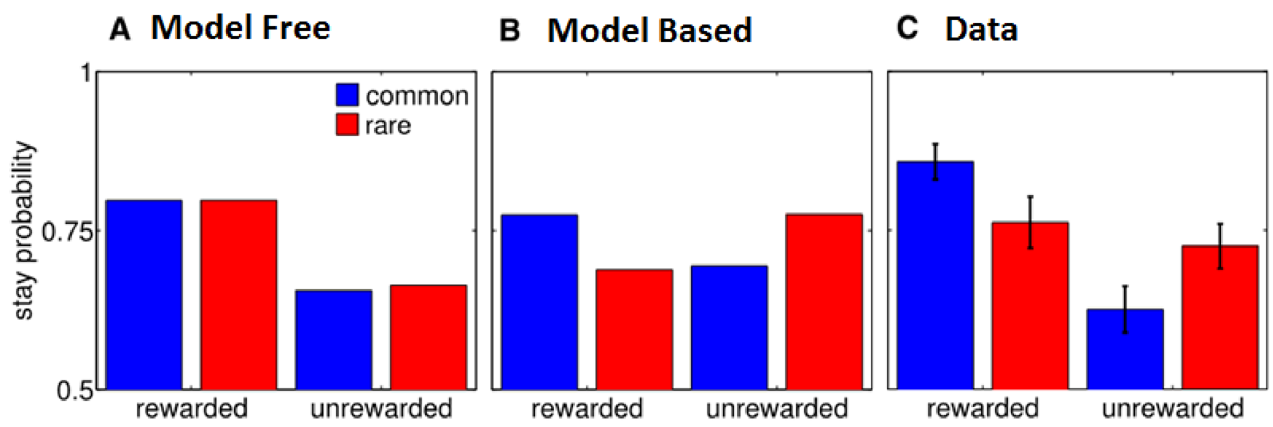


Figure A: Model free agents would be expected to make the same first stage choice after a trial that was rewarded, regardless of whether it entered the second stage via the common or the rare transition path.

Figure B: Model based agents would be expected to take advantage of their knowledge of the environment. If a trial was rewarded after entering the second stage via a common transition path, the agent would be expected to make the same first stage choice again. In contrast, if a trial was rewarded after entering the second stage via a rare transition path, it would be expected to make the opposite first stage choice, in order to maximize the chance of reaching the same second stage state. The same holds true in reverse for unrewarded trials

Figure C: The actual data. Note the results were a mix of both the model free and the model based expectations, suggesting human decision making is a mixture of both processes. (Daw et al., 2011)

Percentile Analysis of Parameter Distribution

	Eta 1st	Eta 2nd	Beta 1st	Beta 2nd	Lambda	W	P	Log Likelihood
25th percentile	0.4483	0.2118	3.298	2.6934	0.4154	0.094	0.0524	167.8667
50th percentile	0.5813	0.42	5.1938	3.6936	0.5595	0.3704	0.1413	197.2618
75th percentile	0.9059	0.7083	7.4349	5.1018	0.9198	0.5221	0.2127	227.5732

Multiple Regression Analysis

	Estimate	Std. Error	t value	Pr(>|t|)	Significance
(Intercept)	583.328	59.109	9.869	8.44E-07	***
w	123.931	60.2	2.059	0.06402	x
Eta 1st	-135.071	57.283	-2.358	0.03795	*
Eta 2nd	185.656	57.789	3.213	0.00827	**
Beta 2nd	16.121	8.103	1.99	0.07208	x

Significance Codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘x’ 0.1 ‘ ’ 1

Correlations between Model Parameters and Reaction Time

	Eta 1st	Eta 2nd	Beta 1st	Beta 2nd	Lambda	W	P	RT 1	RT 2
Eta 1st		0.419	0.084	-0.281	-0.488	0.424	-0.153	-0.167	-0.303
Eta 2nd	0.106		0.05	-0.26	-0.029	0.267	-0.003	0.472	0.082
Beta 1st	0.758	0.854		-0.285	-0.703	0.723	-0.274	0.224	-0.12
Beta 2nd	0.291	0.33	0.284		0.504	-0.413	0.653	0.203	0.357
Lambda	0.055	0.915	**0.002**	0.046		-0.746	0.419	0.164	0.357
W	0.102	0.318	**0.002**	0.111	**0.001**		-0.241	0.242	-0.045
P	0.571	0.991	0.305	**0.006**	0.106	0.368		0.446	0.448
RT 1	0.536	0.065	0.405	0.45	0.544	0.367	0.083		0.403
RT 2	0.254	0.763	0.659	0.175	0.175	0.869	0.082	0.122	

Correlation analyses were conducted between all possible pairwise combinations of the parameters used to define the model, as well as first and second stage average reaction times. The top diagonal lists the calculated correlation coefficients, while the bottom triangle includes the corresponding p-values. Correlations that were significant ($p < 0.01$) have been bolded and highlighted.

Correlation between Model Parameters and RT1, with the inclusion of possible interactions with w

	Estimate	Std. Error	t value	Pr(>|t|)	Signif.
(Intercept)	503.644	205.893	2.446	0.0582	x
w	226.497	488.656	0.464	0.6625	
eta1	-74.744	156.736	-0.477	0.6536	
eta2	156.167	124.282	1.257	0.2644	
beta1	0.775	11.561	0.067	0.9492	
beta2	4.752	16.261	0.292	0.7819	
lambda	99.753	171.241	0.583	0.5855	
p	286.181	230.031	1.244	0.2686	
w:eta1	32.475	397.671	0.082	0.9381	
w:lambda	-153.33	602.135	-0.255	0.8091	
w:p	-473.451	823.391	-0.575	0.5902	

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 'x' 0.1 ' ' 1*

Discussion

- No significant correlation between reaction times and *w*, although a positive trend observed with inclusion of select other parameters
- Many possible reasons. Simple reasons could include subject fatigue or small sample size
- Only looked at subject's data as a whole – did not account for subject's parameters at discrete points throughout the study, or how parameters fluctuated over time
- Could be evidence of neural precompiling and caching
- Follow up studies could include comparisons between reaction time restricted groups (must respond in under ~750 msec, must respond between 1000 and 2000 msec, must respond in over 2500 msec) on similar task, use of cycle detection algorithms to identify subject fatigue, and finer-grained model construction

References

Daw, Nathaniel D., Samuel J. Gershman, Ben Seymour, Peter Dayan, and Raymond J. Dolan. "Model-Based Influences on Humans' Choices and Striatal Prediction Errors." *Neuron* 69.6 (2011): 1204-215.

Solway, Alec M., and Matthew Botvinick. "Goal-Directed Decision Making as Probabilistic Inference: A Computational Framework and Potential Neural Correlates." Psychological Review 119.1 (2012): 120-54.