The General Model for Negative Priming

Hecke Schrobsdorf [1,2], Matthias Ihrke [1,3], Jörg Behrendt [1,4], Marcus Hasselhorn [1,4], J. Michael Herrmann [1,2]


Positive Priming Theories ctd.

Abstract

Negative priming is characterized by longer reaction times when responding to a stimulus which have been actively ignored recently. A central problem of the interpretation of the NP effect is the lack of agreement about the underlying mechanisms. Over the past 20 years, various theoretical accounts have been developed to explain NP. However, empirical evidence does not clearly favour one theory over the others. One of the reasons why it is so hard to decide between the different theoretical accounts, is the lack of a concrete computational framework.

We therefore developed a general model for stimulus-based action selection which attempts to incorporate all mechanisms relevant to selective attention. The concrete implementation incorporates building blocks for feature detection, feature binding, semantic representation, action planning and episodic memory as well as a control unit that keeps track of higher goals. Reaction time differences in various priming conditions emerge by an interplay of all model components.

We present different paradigms of behavioral negative priming experiments and review how the major theories explain decision making and negative priming. After the introduction of the model structure together with its implementation, we point to the behavior of the general model in the different priming conditions and the integration of different theories.

The Psychological Effect

Negative Priming Paradigms

In Negative Priming experiments, subjects have to respond to a target and ignore a distractor. This negative priming was also found by presenting auditory and tactile stimuli.

Concrete Implementation

Feature Layers

Feature activations $c, a, w$ with rise $\tau_{\text{rise}}$ towards input $I$. In the absence of input, they passively decay with $\tau_{\text{decay}}$. Via the bindings activation between features of the same object balance. The feature instance that defines a target as such is amplified with gain $\tau_{\text{gain}}$. With retrieval strength $r$ times strength of the memory trace $m$ on the old episode $a_{\text{old}}$ is retrieved.

$$d \frac{dt}{dt} c = \tau_{\text{rise}}(1 - c) + \tau_{\text{decay}}(a_{\text{old}} - c) \\
+ \sum_{i} \tau_{\text{gain}} \cdot \cdot \cdot \cdot$$

Bind Layer

Bindings are initiated at stimulus onset, their synaptic strength $b$ adapts with time constant $\tau_{\text{bind}}$ towards a maximum strength, if $p = 1$, i.e. the binding is currently perceived. Bindings that share several, but not all features cause a decayed event.

$$d \frac{dt}{dt} b = \tau_{\text{bind}} m \cdot \cdot \cdot \cdot \cdot \cdot$$

Semantic Layer

Input to the semantic layer $S$ from shape $s$ and word $w$ layer with synaptic strengths $\theta_{\text{pair}}(s, c), \theta_{\text{pair}}(w, c)$. Threshold $\theta$ adapts with time constant $\tau_{\text{sem}}$ to an active activity.

$$d \frac{dt}{dt} S = \theta_{\text{pair}}(s, c) \cdot S - \theta_{\text{pair}}(w, c) \cdot \cdot \cdot \cdot \cdot \cdot$$

Action Layer

Depending on the task, the Central Executive chooses a mapping of node and identity of superthreshold semantic representations onto the fixpoints $A$ of the dynamics of the action layer. Between trials, where $l = 0$, actions are initiated by driving the decision not to react $a_{\text{not}}$, which also feeds into this layers adaptive threshold.

$$d \frac{dt}{dt} a = \tau_{\text{sem}} \cdot \cdot \cdot \cdot \cdot \cdot$$

Episodic Memory

After a response, the entire episode, i.e. activation variables of the feature layers that bindings and semantic activations as well as the action values, are stored in episodic memory. The memory trace is assigned an initial strength $m$ that decays with time constant $\tau_{\text{sem}}$. The resulting strength of retrieval $r$ is a normalized sum of the number of matching features $\theta_{\text{pair}}(c, w)$ and correct bindings $r_{\text{correct}}$.

$$r = \tau_{\text{sem}} \cdot \cdot \cdot \cdot \cdot \cdot$$

Imago-Semantic Action Model [10]

The ISAM decides between target and distractor via a threshold that adapts to a global activation level. The target is singled out by a semantic feedback loop. A reaction is triggered as soon as only one object representation is above threshold. A computational implementation with rate equations shows a rich and promising behavior of the ISAM also with single object stimuli.

Distractor Inhibition [3]

In this controversial model, NP is produced by an inertia effect of the underlying artificial neural network. In order to distinguish target from distractor, the network inhibits the distractor. When the input is switched off, the remaining inhibition produces an inhibitory rebound of the distractor. To respond to the former distractor in the probe trial takes longer as the activation has to rise from below baseline.

Episodic Retrieval [7, 2]

The trial onset triggers a retrieval of the former trial from episodic memory. A ‘do not respond’-tag, that was attached to the distractor in the prime episode is retrieved as well. This tag is conflicting with the current task to respond to the former distractor, it has to be removed and thus causes the delay observed in the distractor condition.

Recent advancements state that similarities of prime and probe episode trigger mainly the retrieval of the prime reaction. The more similar the trials are, the stronger is the retrieved representation. In classical negative priming experiments NP-trials always change the reaction whereas positive priming always requires the repetition of reaction times, in classical effects.

Dual Mechanism Hypothesis [5]

Negative Priming might be produced by an interplay of several mechanisms, depending on stimuli, strategies, task etc. Therefore, in settings where memory strategies are an inherent part of the paradigm, memory mechanisms might be more prominent whereas in other situations the inhibition view might account better.

Conclusion

In an exemplary interdisciplinary project that combines computational neuroscience and cognitive psychology, we interactively developed elaborated experimental paradigms together with the aid of a comprehensive model.

The model incorporates various components: perception, attention, memory, semantic representations and action selection, thereby allowing us to cross test concuring theories for negative priming.

Reaction time differences in negative priming conditions emerge due to an interplay of all model components.

A model based analysis of EEG data eases the finding and interpretation of neural correlates of negative priming.

Outlook

In order to cross test the different theories, parameters will be collated to single approaches to tune the impact of a single theory on the model behavior.

Peculiarities of certain features have to be translated into topologies of the specific feature layer to apply the model to a broader range of stimuli and paradigms.

Integrating effects of cognitive aging into the model.

Application of the general model to control autonomous robots.

References


