Hypothesis

Shiro Takagi

2021/4/28

1 Preface

1.1 Problem setting

We consider the distributed representation. A memory of experience $\boldsymbol{x} \in X \subset \mathbb{R}^n$ is a function $f : X \mapsto \Theta$. For simplicity, we denote a memory by $\boldsymbol{\theta} \in \Theta \subset \mathbb{R}^p$. Note that n and p are dimension of vector $\boldsymbol{x}$ and $\boldsymbol{\theta}$.

We discuss short term memory and long term memory. Short term memory is a function $f^s : X \mapsto \Theta^s$, where $\boldsymbol{\theta}^s \in \Theta^s \subset \mathbb{R}^{p^s}$. Long term memory is a compositional function $f^l \coloneqq f^c \circ f^s$, where $f^c : \Theta^s \mapsto \Theta^l$ is a memory consolidation function and $\boldsymbol{\theta}^l \in \Theta^l \subset \mathbb{R}^{p^l}$. Also, we consider memory retrieval function $f^r : \Theta^l \mapsto \Theta^s$. We assume that p^l is finite and fixed

1.2 Goal

Our goal is to find optimal memory consolidation function f^c and memory retrieval function f^r, given some criterion. Followings are what we think are desirable properties for these functions to have:

- long term memory retains short term memory's information as much as possible
- long term memory is retrievable by memory retrieval function
- memory retrieval function retrieves stored information as much as possible

Another thing to consider is how to combine long term memory to short term memory. Humans seem to elegantly and naturally exploit these two memory to do a task. Thus, long term memory should be encoded and decoded such that it can be exploited easily.

1.3 Memory representation

We represent memory as just a vector with no structure $\boldsymbol{\theta}$. However, the relation between each component of a distributed representation are generally asymmetric. For example, parameters of fully connected neural network construct a hierarchical structure. Therefore, considering an optimal structure of a distributed representation, given some criterion, is another issue to consider.

Bunch of studies discussed how to construct parameters for to do tasks well (short term memory). My aim is to elucidate an optimal structure for long term memory.

References