The Algebra of Cognitive States: Towards modelling the Serial Position Curve

Stefan Reimann (stefan.reimann@protonmail.com)
Department of Psychology, Cognitive Psychology
University of Zurich, Switzerland

Abstract

A computational framework for modelling storage and retrieval of information in human working memory is proposed. The aim is to analyse the corresponding algebra alone, especially with regard to its congruence with empirical findings including the serial position curve. That algebra builds on the high-dimensional holographic representation of information together with two operations for computation: multiplication for binding and addition for bundling. Unlike other models, the bundling operation defined is not associative and preserves serial order information in terms of activation gradients. Consequently, the cognitive states representing a memorised list exhibit a primacy as well as a recency effect generically. The typical concave-up and asymmetrically shaped serial position curve is derived as a linear combination of those gradients. The serial position curve for cued recall, including similar items, is derived within this formalism. Quantitative implications of the algebra are shown to agree well with empirical data from basic cognitive tasks.

Keywords: human working memory; activation gradients; serial position curve; holographic representation; high-dimensional computing

Introduction

Human Working Memory is commonly regarded as a functional subsystem of memory, whose goal is to hold and to organise information for some short period of time in order to make it available for higher cognitive processes (Cowan, 2017). Experiments in this field rely on the subtle construction of input data such as memory lists and produce output data such as recall probabilities or response times (Murdock, 1974; Kahana, 2012; Oberauer et al., 2018). Among these, the most prominent finding is the serial-position curve, which shows the accuracy of item retrieval varying as a function of serial position in a memory list, averaged over a sample of participants. As observed across (probably all) immediate memory tasks, it has a concave-up shape and is asymmetric. Its particular shape depends on the particular cognitive task. For example in recognition and in cued (probed) recall the serial position curve shows a strong recency effect, while the primacy effect is weak. Strong primacy effects are seen in forward recall, while recency effects are strong in backward or in free recall.

To describe particular aspects of the functioning of the human working memory, models with different characteristics have been used, differing both in terms of the medium in which the information is stored and the storage operations used. Models include local code models such as REM
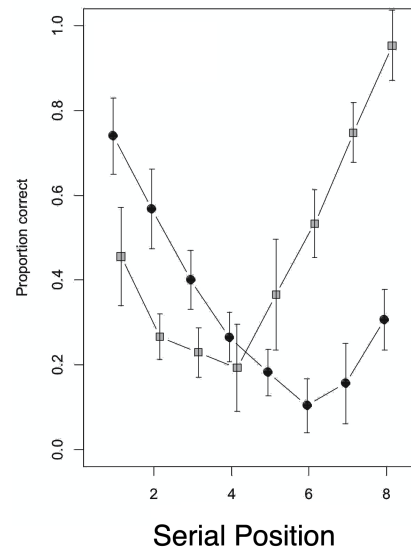

Figure 1: **Primacy effect and recency effects**: Data are for immediate forward and backward serial recall (Oberauer et al., 2018), but are similarly in other immediate tasks (Murdock, 1974; Kahana, 2012)

(Shiffrin & Steyvers, 1997), distributed models of memory such as SOB (Farrell & Lewandowsky, 2002) and TCM (Howard & Kahana, 2002), as well as holographic models such as TODAM (Murdock, 1982, 1993) which uses high-dimensional probabilistic encoding for the holographic representation of information (Plate, 1991). Holographic models gain from the properties, which are implied by high dimensionality together with randomness, see (Kanerva, 2009) for an overview about the framework of high-dimensional computing.

The holographic approach appears as a natural candidate to model the functioning of cognitive processes. Input items evoke activity patterns in the respective neural field; The fact that these representations are sparse and the consequences thereof are not explicitly considered in this note. Computation consists in transforming those patterns according to two elementary operations: The additive-like superposition realises the bundling of item information, while multiplication realises binding of items. The high-dimensional space

of binary patterns together with these two operations form a high-dimensional algebra governing storage and computation in this system.

Before giving an outline of the paper, a remark seems worthwhile: The aim is not to provide a full-blown model rather than to propose an elementary computational structure, an algebra, on top of which a model could be constructed. The main question is, how much of experimental findings can already been described on the basis of that algebra alone.

The outline of this paper is briefly as follows: Firstly, the state-space is defined as a high-dimensional Hamming space (eq 1) equipped with some distance on it. A similarity measure is proposed which is derived from that distance. It allows both, to judge about the familiarity of two states as well as about their distinctiveness (eq 2). Computing is by manipulating states according to two operations on that space: multiplication for binding and a not associative addition for bundling. This completes the definition of the algebra (eq 4) to be considered. Non-associativity is an essential feature of that bundling since it implies that the sum of components depends on their sequential ordering (eq 5). As a consequence, information about the order of sequentially presented list items is conserved. The corresponding left-associative sum and the right- associative sum of list items correspond to states exhibiting a recency and a primacy gradient, respectively (Fig 4). As applications basic cognitive tasks such as item recognition and probed recall are considered. The typical concave-up and asymmetrical shape of the serial-position curve is derived as a mixture of these two activation gradients (Fig 9).

The algebra of cognitive states $(\mathbb{X}, +_p, *)$

The state-space   In the course of perceiving a physical item, the corresponding sensory input invokes an activity pattern in the neuronal field it is projected to. That way, each physical item can be represented by a binary pattern, in which $1's$ indicate active neurons, while $0's$ indicate inactive ones. Due to the size and structural complexity of the neuronal correlate, patterns are described by high-dimensional random binary vectors. These patterns are the states of the cognitive system. The state-space therefore is

$$\mathbb{X} = \left(\mathbb{X}_q^N, d\right). \tag{1}$$

$N > 100$ is its dimension, q is the degree of sparseness, i.e. the mean activity of a state, and d is some metric on $\mathbb{X}_q^N$.

The state-space is a (metric) Hamming space allowing for some similarity measure derived from the distance d. This measure should respect both: the closeness of two states as well as their distinctiveness as points in the state-space. A cosine-similarity only reveals information about closeness since it is locally defined. In a probabilistic setting, two points are the more difficult to distinguish, the less likely it is to find another state at random which is 'in between' the two. To capture this, the definition of similarity must contain global information about the state space.

Definition (Similarity). *The similarity of two states having distance d from each other is*

$$S(d) := e^{-\kappa F_{\mathbb{X}}(d)}, \qquad \kappa > 1 \tag{2}$$

where $F_{\mathbb{X}}(d) = \mathbb{P}_{\mathbb{X}}[D \le d]$ is the distribution function for distances on $\mathbb{X}$.

Different items are represented by uncorrelated states, while similar items will be represented by similar states. $\kappa > 1$ is chosen to have highest sensitivity with respect to almost identical or near-by states.

The operations   The two operations to be defined on the state space correspond to binding and bundling. Two items are (associatively) bound to each other, if one can be retrieved by cueing with the other item. The corresponding formal operation is *multiplication* $*$, which is defined in eq 3. Binding of items happens by simultaneously activated components in the neural pattern. This similarity measure directly relates to a recall probability or accuracy of retrieval.

Bundling means collecting items by adding their respective states. Assume that two neurons X and Y converge on a third neuron Z. If both are inactive, i.e. $x = y = 0$, neuron Z will also be, $z = 0$, while if both are active, Z will be active, i.e. $1 + 1 = 1$. If only X or Y is active, it depends on some threshold, whether Z is active. If the activation threshold is low, $p \approx 0$, Z is likely to be active, while if if the activation threshold is high, $p \approx 1$, Z will remain inactive. Addition x_p is defined in eq 3.

$*$	0	1		$+_p$	0	1
0	1	0		0	0	ζ
1	0	1		1	ζ	1

$$\tag{3}$$

where $\zeta \in \{0, +1\}$ is random with $\mathbb{P}[\zeta = 0] = p$. This completes the definition of the algebra used to calculating with cognitive states.

$$\left(\mathbb{X}, +_p, *\right) \tag{4}$$

In the following, its elementary properties are further investigated. What properties are already implied by this elementary algebra and how much of empirical findings can be already described by those?

Bundling preserves sequential information in the memory list

Usually, bundling is realised by vector-addition (Schlegel, Neubert, & Protzel, 2020), which is commutative and associative, so that $x + (y + z) = (x + y) + z = z + (y + z)$, i.e. the order of components doesn't matter. That is: If addition is associative, sequential order information is lost!

Observation. *For $0 < p < 1$, addition $+_p$ is not associative.*

$$x +_p (y +_p z) \ne z +_p (y +_p x) \tag{5}$$

Note that, if $p = 1$, addition equals component-wise *AND*, while for $p = 0$, addition is component-wise *OR*. These operations are associative.

In the following, the state resulting from left-associative addition is denoted by **L**, i.e. $\mathbf{L} = (x +_p y) + z$, while the state resulting from right-associative addition is denoted by $\mathbf{R} = x +_p (y +_p z)$. For the sake for readability, I will write $+ = +_p$ in the following, while assuming that $p = \frac{1}{2}$.
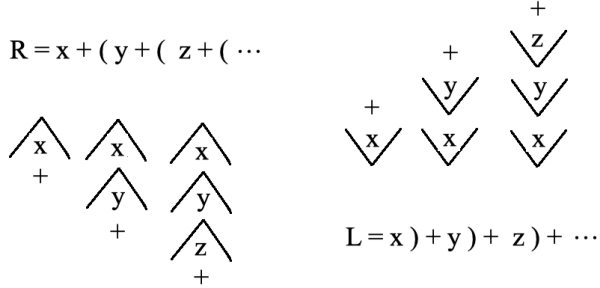
$$R = x + (y + (z + (\cdots$$

$$L = x\,) + y\,) + z\,) + \cdots$$

Figure 2: In right-associative **R** addition, early items are kept prominent, while in left-associative **L** addition, later items are superposed on earlier ones.

The states representing a memory list   **L** and **R** states can be constructed for a list of any length. Construction starts from a pre-experimental state η and proceeds by iteratively adding items to the memory states **L** and **R** according to left-associative addition and right-associative addition to the respective branch as follows: For the **L**-state

$$
\begin{aligned}
\mathbf{L}^0 &= \eta \\
\mathbf{L}^a &= \eta + a) \\
\mathbf{L}^b &= \eta + a) + b) \\
&\vdots \\
\mathbf{L}^\Lambda &= \Big((((\eta + a) + b) + c) + \ldots) + f) + g \Big),
\end{aligned}
$$

while for the **R**-state

$$
\begin{aligned}
\mathbf{R}^0 &= \eta \\
\mathbf{R}^a &= \eta + (a \\
\mathbf{R}^b &= \eta + (a + (b \\
&\vdots \\
\mathbf{R}^\Lambda &= \Big(\eta + \big(a + (b + (c + (\ldots + (f + g)))) \big) \Big)
\end{aligned}
$$

After its sequential presentation, the memory list $\Lambda = (A, B, C, \ldots)$ is thus represented by the two states

$$\mathbf{L} = \Big((((\eta + a) + b) + c) + d) + f) + g \Big) \quad (6)$$

$$\mathbf{R} = \Big(\eta + \big(a + (b + (c + (d + (f + g)))) \big) \Big), \quad (7)$$

In (Murdock, 1982) η is assumed to be empty, while in (Franklin & Mewhort, 2015) it comprises a holographic collection of items and item-item associations. $a, b, c, \ldots$ are the cognitive states representing the physical list items $A, B, C, \ldots$. These states preserve the serial order of items in

the memory list in that distances change monotonously along subsequent items, see Fig. 3

$$
\begin{aligned}
d(\eta, \mathbf{R}) &< d(\eta, \mathbf{L}) & (8) \\
d(a, \mathbf{L}) &> d(b, \mathbf{L}) > d(c, \mathbf{L}) > \ldots & (9) \\
d(a, \mathbf{R}) &< d(b, \mathbf{R}) < d(c, \mathbf{R}) < \ldots & (10)
\end{aligned}
$$

Correspondingly, both states inherit serial order in that item distances increase along **R**, while they decrease along **L**, see Fig 3. These distance gradients directly translate into activation gradients.
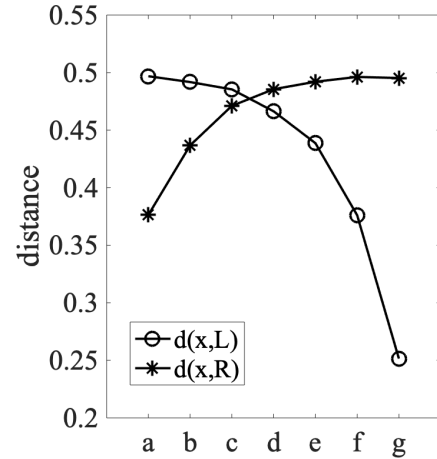

Figure 3: **Distance profiles** of the two states **L** and **R** as in eq 8 ff. **L** has smallest distances to the most recent items, while **R** is closest to the early list items.

Implied activity gradients

From the concept of similarity, two other concepts can be immediately derived: activation and memory strength. The intuition is closely related to the idea of a projection. Given that the memory state **M** represents a memorised list, and that a cue item is presented. The cue item activates the memory state more, the more similar it is to that memory state (Hintzman, 1984). Conversely, the more the corresponding memory element is engraved in the memory state, the more the memory state is activated by the cue state.

Definition (Activation). *Let **M** be a memory state constructed during representing some memory list. A cue state x activates the memory state **M** according to their similarity, see eq 2*

$$\alpha_{\mathbf{M}}(x) := S\big(d(x, \mathbf{M})\big). \quad (11)$$

*The activation gradient of **M** is the vector $\alpha_{\mathbf{M}}$ with components $\alpha_{\mathbf{M}}(x)$, where x is a state representing a list item.*

In terms of strength theory, $\alpha_{\mathbf{M}}(x)$ is the strength by which x is memorised in **M**. One might also call $\alpha_{\mathbf{M}}(x)$ the familiarity of x given **M**.

Consequently, the distance gradients in eq 8 ff directly translate into activity gradients, see Fig. 4. Since activation
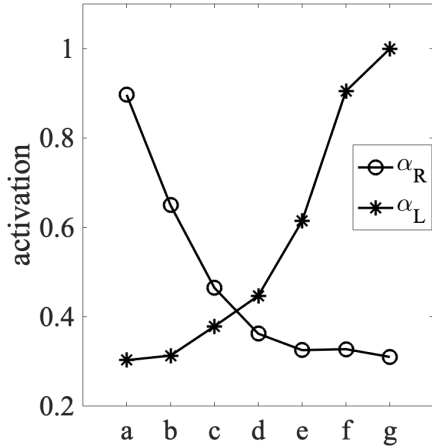
Figure 4: **Primacy and recency gradients** implied by the two states **L** and **R** are $\alpha_{\mathbf{R}}$ and $\alpha_{\mathbf{L}}$.

Experimental data indicate that the recency effect does not depend on list length and shows a slightly sigmoid curve shape, see Fig 5 (left). Both empirical observations are well captured by the modelling algebra proposed, see Fig 5 (right).
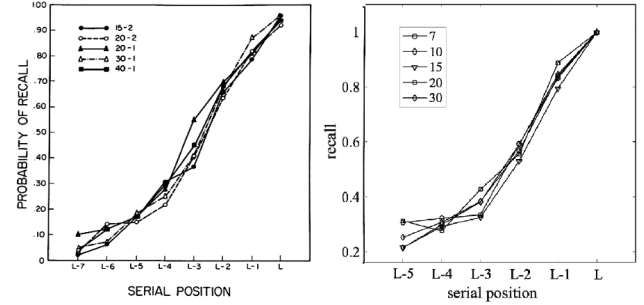

Figure 5: The recency effect does not depend on list length. Left: Experimental data from Murdock (Murdock, 1982), Right: Simulated data from the model for various list length'.

as well as strength are increasing functions of similarity and hence decreasing functions of distance, **L** implies a *recency gradient* $\alpha_{\mathbf{L}}$, while **R** implies a *primacy gradient* $\alpha_{\mathbf{R}}$.

Activation gradients are nowadays widely accepted to play an important role in working memory. Various mechanisms have been discussed as sources of these gradients, see (Oberauer, 2003). In many models including TODAM, TCM and SOB, these gradients are separately modelled and superimposed on top of the model. In contrast, these gradients directly result from the bundling operation defined in eq 4 and its non-associativity: While non-associativity preserves information about serial order, right-associative addition and left-associative addition imply the primacy and the recency gradient, respectively.

The response function for recognition and recall

After presentation of a memory list, the participant has to fulfil some task. Most cognitive tasks involve cues such as cued item recognition or cued recall, associative or serial. The answer the participant gives is the result of a decision process which depends on both, the memory state as well as the cue. The response function in recognition only depends on familiarity, while the response function in recall additionally depends on distinctiveness (Murdock, 1982). Thus it is reasonable to make the response function a function of activation as defined in eq 11.

Definition (Response function). *The response function given a cue x facing the memory state* **M** *is an increasing function of induced similarity, e.g.*

$$\Phi(x \,|\, \mathbf{M}) = \alpha_{\mathbf{M}}(x) \qquad (12)$$

Accordingly an activation gradient directly translates into a serial position curve. Particularly, the recency effect refers to the activation gradient of the **L**-state, while the primacy effect corresponds to the activation gradient of the **R**-state.

Application to some basic cognitive tasks

In this section some examples are presented to demonstrate how the formalism works, i.e. how to describe tasks such as cued recall in this formalism. Results are direct consequences of the algebra defined, i.e. no further assumptions are made. In the following only the **R**-state is concerned, i.e. states are bundled according to right-associative addition, while corresponding brackets are skipped for the sake of readability.

Repetition increases strength

It is intuitively expected that a repeated occurrence of an element in a list will increase its coding strength. This effect is indeed observed in the model. As a benchmark, consider the list $\Lambda = (A, B, C, D, \ldots)$, in which all items are different. In $\Lambda^{(1)}$ a neighbouring pair is similar, e.g., $B \sim C$. In $\Lambda^{(2)}$, $B \sim D$ and so forth. k can be regarded as the lag from B until the similar item. Fig 6 shows the serial position curves for lists Λ, $\Lambda^{(1)}$, and $\Lambda^{(2)}$. Note that the coding strength of B is increased by any other item which is similar to B, while the strengthening is greater, the smaller the lag is, i.e. the effect of $C \sim B$ on the coding strength of B is larger than the effect of $D \sim B$ on the coding strength of B.

Cued recall

Cued associative recall In this task, the participant is presented a paired memory list $\left(A - X, B - Y, C - Z, \ldots\right)$. After memorizing this list, a memory item, i.e. a member of some pair, is presented as a cue, and the participant is asked to identify the memory item, which was bound to that cue item. The memory state corresponding to the paired list is

$$\mathbf{R} = \eta + a*x + b*y + c*z + \ldots, \qquad (13)$$

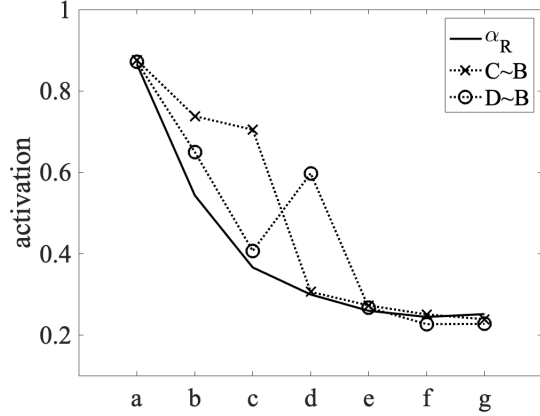where $a*x$ is the state representing the binding between items X and A in the list.

Figure 6: **The effect of similar items on activation** The solid black curve is the activation profile of state **R** for the list Λ in which all items are different. Doted lines are the profiles if that list contains one item, e.g. C or D, which is similar to item B.

When a memory item X is presented as a cue and the task is to retrieve the item which is bound to X in the list, consider the activation of

$$x * \mathbf{R} = x * \eta + a + a*b*y + a*c*z + \ldots \quad (14)$$

The activation $\alpha_{x*\mathbf{R}}$ attains its maximal value for $\alpha_{x*\mathbf{R}}(a)$, see Fig. 7. Thus the cue X activates the A component most, so that the participant will answer *" X is bound to A."* , with some probability. Analogously, if the cue is Y, the activation $\alpha_{y*\mathbf{R}}$ attains it maximum in B, so that B is retrieved, and so forth. These maximal points form a curve, which is identical to the activation gradient $\alpha_{\mathbf{R}}$.

Retrieval from similar contexts Assume that the paired list $\left(A-X, B-Y, C-\tilde{X}, D-Z, \ldots\right)$ is given, in which items A and C are bound to similar contexts X and $\tilde{X}$. The corresponding state yields

$$\mathbf{R} = \eta + a*x + b*y + c*\tilde{x} + \ldots \quad (15)$$

Cueing with X will not only retrieve A but also C, just to a lesser extend. The effect of cueing with x is displayed when considering the activation gradient $\alpha_{x*\mathbf{R}}$, see Fig. 8. The gradient has two peaks, one at a and a weaker one at c, saying that cueing with X reveals two items, A and C. Cueing with Y uncovers only one, which is B.

In the recall task, the participant has to make a choice between the two alternative items bound to X. Thus invoking Luce's choice axiom, the probability to recall X yields

$$P(a|x) = \frac{\alpha_{x*\mathbf{R}}(a)}{\alpha_{x*\mathbf{R}}(a) + \alpha_{x*\mathbf{R}}(c)} \quad (16)$$

which is less than the probability to recall a without an alternative. The existence of an item similar to the cue impairs the corresponding recall.
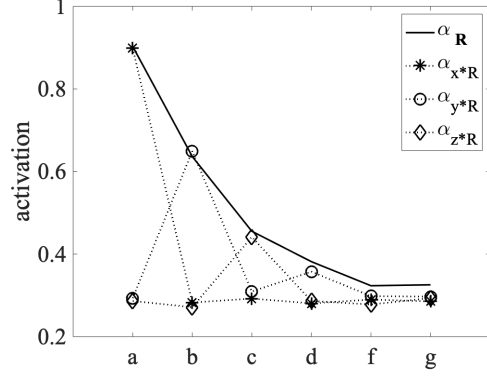

Figure 7: **Cued Recall**: Given cues such as x, y, z, the corresponding activations $\alpha_{x*\mathbf{R}}, \alpha_{y*\mathbf{R}}, \alpha_{z*\mathbf{R}}$ are plotted, see eq 14. The cue x causes the activation $\alpha_{x*\mathbf{R}}$ to have a peak at the corresponding item, which is a. $\alpha_{\mathbf{R}}$ is the activation profile of the **R**-state.
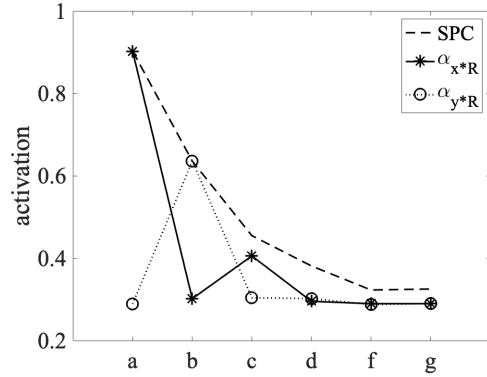

Figure 8: **Recall of items**: Activation profiles of $\alpha_{x*\mathbf{R}}$ ($*$), and $\alpha_{y*\mathbf{R}}$ ($\circ$). The activation profile $\alpha_{\mathbf{R}}$ of the **R** state by distinct list items is shown as a reference.

Putting things together: The serial position curve

During memorizing a list, the two states **R** and **L** are constructed. Since there is no a priori reason to favour one over the other, I assume that both cognitive states **L** and **R** coexist and are the components of a *memory state* **M**,

$$\mathbf{M} = \begin{pmatrix} \mathbf{L} \\ \mathbf{R} \end{pmatrix}. \quad (17)$$

A single cue thus activates both components. The total activation of the memory state **M** is a linear combination of the activation gradients of its two components.

$$\alpha_{\mathbf{M}} = \rho\,\mathbf{R} + \ell\,\mathbf{L} \quad (18)$$

where ρ and ℓ are non-negative parameters governing the mixture of respective activations. The response function to a cue is $\Phi(x|\alpha_{\mathbf{M}})$, so that the serial position curve is the graph $\Phi(x|\alpha_{\mathbf{M}})$, where x is a state representing a list item, see Fig. 9.
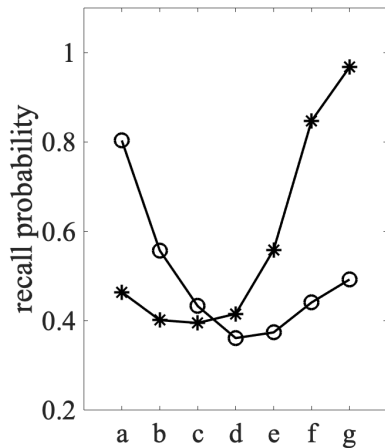
Figure 9: **The Serial Position Curve** is simulated for different pairs of parameters according to eq. 18. It shows a strong recency effect and a weak primacy effect for $\rho = 0.4, \lambda = 0.9$, while for $\rho = 0.9, \lambda = 0.4$ there is a strong primacy effect and a weak recency effect.

The serial position curve thus results from the linear combination of the primacy gradient $\alpha_\mathbf{R}$ and the recency gradient $\alpha_\mathbf{L}$. As seen in Fig 9, a large ρ together with a small ℓ makes the recency effect, while a small ρ together with a large ℓ leads to a prominent primacy effect. The relative strength of the primacy and the recency effect will generally depend on the experimental set-up, including the task to be performed. For example in recognition and in cued (probed) recall the serial position curve shows a strong recency effect, while the primacy effect is weak. Strong primacy effects are seen in forward recall, while recency effects are strong in backward or in free recall.

Conclusion and out-look

In the previous sections, an elementary algebra (eq. 4) for storage and retrieval of information in basic cognitive tasks was proposed. The aim was not to present a full-blown model but to investigate how far one can get with the algebra alone.

Item information and associative information are represented by two operations, bundling and binding, respectively. If bundling is realised by an associative operation such as ordinary (vector-) addition, information about sequential order is lost. On the other hand, tasks such as serial recall require that order information. Consequently in corresponding models order information has be has to be implemented separately. This can be achieved by postulating serial position markers, chaining by associative mechanisms between consecutive items, or weight functions varying over serial position governing the recency and the primacy effect.

This is different in the approach presented: Information about sequential ordering is preserved. This is due to the non-associativity of the addition operation by which item in-

formation is bundled into a memory state. Reading from that state thus reveals order information necessary to related tasks, which is represented by corresponding gradients. Activation gradients are implied rather than postulated separately. The serial position curve comes as a linear combination of both. Its shape is concave-up and asymmetric as observed as a typical experimental finding, see Fig 1 for experimental data and Fig 9 for simulations of our model.

As already mentioned, the aim was not to present a full-blown model but to investigate how far one can get with the algebra alone. So it does not come as a surprise that several experimental observations were not captured. For example, while the recency effect does not depend on list length, the primacy effect does. This robust finding cannot be explained by our algebra alone but needs an additional assumption about attention, which then imposes an additional constraint on den attention gradient. Furthermore, serial recall can not be described by our algebra alone but needs an additional assumption such as output-suppression, as supposed in many models, or an other feedback mechanism, see (Franklin & Mewhort, 2015).

The cognitive algebra proposed appears to provide a reasonable basis for modelling since it generically implies several features that fit empirical observations quite well, in a qualitative sense in that no attempt was made to fit data. Modelling then could consist in carefully adding assumptions on top of the cognitive algebra such as discussed above.

Acknowledgement

References

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic bulletin & review*, *24*(4), 1158–1170.

Farrell, S., & Lewandowsky, S. (2002). An endogenous distributed model of ordering in serial recall. *Psychonomic bulletin & review*, *9*(1), 59–79.

Franklin, D. R., & Mewhort, D. (2015). Memory as a hologram: An analysis of learning and recall. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *69*(1), 115.

Hintzman, D. L. (1984). Minerva 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*(2), 96–101.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299.

Kahana, M. J. (2012). *Foundations of human memory*. OUP USA.

Kanerva, P. (2009). Hyperdimensional computing: An introduction to computing in distributed representation with high-dimensional random vectors. *Cognitive computation*, *1*(2), 139–159.

Murdock, B. B. (1974). *Human memory: Theory and data.* Lawrence Erlbaum.

Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, *89*(6), 609.

Murdock, B. B. (1993). Todam2: a model for the storage and retrieval of item, associative, and serial-order information. *Psychological review*, *100*(2), 183.

Oberauer, K. (2003). Understanding serial position curves in short-term recognition and recall. *Journal of Memory and Language*, *49*(4), 469–483.

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D., Conway, A., Cowan, N., . . . others (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*(9), 885.

Plate, T. (1991). Holographic reduced representations: Convolution algebra for compositional distributed representations. In *Ijcai* (pp. 30–35).

Schlegel, K., Neubert, P., & Protzel, P. (2020). A comparison of vector symbolic architectures. *arXiv preprint arXiv:2001.11797*.

Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: Rem—retrieving effectively from memory. *Psychonomic bulletin & review*, *4*(2), 145–166.