The Cerebellum according to the Ouroboros Model, the “Interpolator Hypothesis”

Knud Thomsen

Research with Neutrons and Muons, Paul ScherrerInstitut, Villigen PSI CH-5232, Switzerland

Received: February 22, 2014 / Accepted: March 12, 2014 / Published: March 31, 2014.

Abstract: The Ouroboros Model offers a novel cognitive architecture with an algorithmic backbone of iterative and self-referential processing. All memory content is organized into meaningful pieces of data, i.e., chunks and schemata, which are laid down as a kind of snapshots of all activations at a relevant point in time. This entails a grainy structure of memory content. Whereas a core process of “consumption” analysis can naturally be defined taking advantage of this parcelisation, it necessitates interpolation for fine nuances, especially for the representation of intermediate values during transients. It is argued that, in the vertebrate brain, essential interpolation functionality is provided by the cerebellum. Findings concerning movement control and perception as well as the involvement of the cerebellum in more abstract, cognitive, tasks can be interpreted as reflecting a function of the cerebellum as a co-processor, i.e., a general-purpose interpolator. The cerebellum is thus boosting representations in cerebral cortex areas, which are reciprocally connected with cerebellar cortex areas. In this paper, it is sketched how the new “interpolator hypothesis” can explain manifold observations, effectively subsuming previous theories of the function of the cerebellum. Directly relating to the BICA challenge, a general-purpose interpolator is postulated as an efficient device providing fine-grained resolution for the representation of diverse content, potentially useful in any cognitive architecture.

Keywords: Schemata, grainy memory structure, interpolation, timing, extrapolation.

1. Introduction

Strikingly conflicting proposals have been made concerning the function of the cerebellum, and, so far, no full consensus has been reached. Here, a novel functional account is presented, which addresses a general need for interpolation and which allows reconciling many of the previous conceptualizations.

The paper is structured as follows. After a summary of key tenets of the Ouroboros Model, the motivation for an interpolation function and the “interpolator hypothesis” are introduced. In terms of implementation, evidence from actual brains is presented, and the relation to existing accounts is clarified. Particular examples supporting the interpolator hypothesis relating to motor control are described, and special cases are discussed demonstrating wide-reaching implications. Conclusions, in particular pertaining to the usefulness of a general interpolator function beyond the Ouroboros Model, end the paper. An early version has been presented before [1].

The aim of this short paper is not to present a full-fledged theoretical model with quantitative details but rather to outline a new overarching perspective suggesting a fresh comprehensive understanding of cerebellar function, which is considered relevant to the BICA challenge, i.e., developing a computational equivalent to the human mind [2].

2. The Ouroboros Model in Short

In a series of recent papers, the Ouroboros Model has been introduced as a novel attempt at explaining a wide range of findings pertaining to cognition and consciousness of natural and also for artificial agents [3-9]. It has been described how, within a single
approach centered around a principal algorithmic process on a suitably structured memory, one can explain human cognitive performance and also formulate prescriptions of how to achieve comparable capabilities with artificial agents implemented in hardware or software, all iteratively and recursively following a similar self-steered evolutionary program.

Minds are seen as primarily data processing entities; an iterative and self-referential universal algorithmic layout on the basis of suitably stored data structures is held as essential [3,5,9].

2.1 Action and Memory Structure

The Ouroboros Model takes memory entries as effectively organized into (non-strict) hierarchies of schemata. Such schematic structures have been described by Immanuel Kant [10], elaborated by Otto Selz[11,12] and also Frederic Bartlett [13]. Memory is made up of meaningful chunks, i.e., combinations of features and concepts belonging together[4]. In living brains, neural assemblies are permanently linked together when once co-activated in the right manner. Later re-activation of any one of the linked features excites an entire schema. In particular, also currently missing features are biased and thus expectations of some type are triggered.

2.2 Principal Algorithmic Backbone

At the core of the Ouroboros Model lies, a self-referential recursive process with alternating phases of data-acquisition and data-evaluation. A monitor process termed “consumption analysis” is checking how well expectations, which are triggered at one point in time, fit with successive activations; these principal stages are identified:

• anticipation,
• Action/perception,
• Evaluation,
• Anticipation.

These steps are concatenated into a full repeating circle, and the activity continues at its former end, like the old alchemists’ tail-devouring serpent called the Ouroboros. Most importantly, there is no detrimental circularity involved if the succession of the processing steps in time is well taken into account: teeth and tail of the name-giving snake belong to well distinct points in time. The outputs of one iteration cycle lay the basis for the next.

Although connections are marked “excitatory” and “inhibitory”, no direct correspondence to nervous structures is intended in this case; “excitatory” simply stands for a link activating the following, receiving entity, and “inhibitory” means that arriving activation dampens or prohibits activity of the recipient. Reprinted with permission from Ref. [3].

The Ouroboros Model can be seen as an extension of the perceptual cycle as originally proposed by Ulric Neisser[14]. The main addition consists in the tight interweaving of a foundation of already existing schemata with perception, triggered by the input at a point in time and the subsequent unfolding of action according to a simple general algorithm with “consumption analysis” at its core. Fig. 1 gives an outline of the processing steps; a more detailed account specifying the genesis and possible extensions of the basic scheme can be found elsewhere [3].

The Ouroboros Model certainly is not the only conceptualization of mental processes, which comprises some processing loop. As distinguishing feature, the emphasis of the crucial importance of components and constituents of concepts, “schemata”, can be seen, as well as the clearly specified core of the processing, i.e., consumption analysis, in which anticipations are set in relation to actually available data. It has been sketched how the Ouroboros Model can be conceptualized as an implementation of Bayesian processing [5,15]; in this view, priors are contributed by the previously established concepts, and hypotheses are generated in this context relating to newly arriving data. Cognitive architectures, which rely on productions, are understood as focusing on one facet of the use of schemata, in this sense employing a
special case of processing contained in the Ouroboros Model [9,16-18]. Any partly activated schema with an open slot acts exactly like a production rule with preconditions specified and conclusion still open. In nice resemblance with processing in the Ouroboros Model, where unity of action (and consciousness) stems from tying all concurrent activity regularly together for consumption analysis, in ACT-R (Adaptive Control of Thought-Rational) only one single production rule is active at any given point in time [19].

From a wider perspective, the “OODA loop”, which comprises the stations: Observe-Orient-Decide-Act, can also be seen as a related theory [20]. It is held that by comparison the Ouroboros Model in terms of consumption analysis does offer an algorithmically much clearer account than merely referring to “feedback”. The cycles of constant quality-control and quality-improvement, “PDCA cycles”, which consists of the stages: Plan-Do-Check-Act, fit into the same mold [21]. Repeated refinement and continuous improvement (“kaizen”) match perfectly with the iterations of the Ouroboros Model in periods or settings without major resets. The results achieved in the frameworks of the OODA-loop and the PDCA-loop for competitive settings nicely complement currently ongoing work on (collaborative) communication [22]. Successful dynamic action in all contexts relies on shared grounding and (conceptual) building blocks as well as on taking everything relevant, in particular, the opponents/partners, fully into account. A difference in sign depends on whether the goal of the involved agents is the establishing or the disruption of cycles of mutual anticipation, overall agreement, understanding and consistency.

It could be claimed that already Otto Selz[11,23] has outlined the principles, which are illustrated in Fig. 1. He was probably among the first to note that a
complete schema can be activated already by a part of its constituents with anticipations pointing to some missing constituents. Schemata are stacked and nested, i.e., structures at one level can function as components for a next higher organizational unit. Anticipations are at the core of the orderly problem solving process; errors occur in cases when only part of the relevant information is effectively used.

A most important points is that the basic principles of memory organization as well as processing stay the same, independent of the content of the linked attributes, which can also stem from very different spheres, e.g., varying in their grounding. Any activation can be incorporated into new entities and contexts; the resulting hierarchy of concepts will not be a strict one, and still, rather well distinct building blocks will be accrued. In particular, self-monitoring does not require any completely different ingredients for meta-cognition: the very same processing steps working with just another set of concepts added, i.e., referring self-reflectively to the actor, do the trick [6]. The claimed benefits of a meta-cognitive loop, “MCL”, for the robustness of intelligent systems are taken to stay the same [24]. A repeated sequence of “Note-Assess-Guide” is a practical shortcut describing what the Ouroboros Model all the time does; powerful general-purpose anomaly-handling strategies are nothing but particular schemata, easy learnable (as abstractions) from experience or instruction.

2.3 Consumption Analysis

As an activation occurs, such as triggered by a sensory precept, associated schemata are excited. Schemata are searched in parallel; the one with the strongest bottom up activation sharing similar features is activated first; see Fig. 1. Other schemata, which possibly are also applicable, are inhibited and their activation is suppressed. With the excitation of a schema, all its constitutive components and linked features are marked; expectations for not yet available attributes are thus triggered. Taking the first selected schema and ensuing anticipations, which are active at that time as reference and basis, consumption analysis checks how successive activations fit into this activated frame structure, i.e., how well lower level input data are “consumed” by the chosen schema. Features are assigned, attributes are “explained away” and inhibited for immediate reuse [25]. In order to avoid preterm fixation and getting bogged down in detail prematurely, it is wise to check first for plausibility on a coarse scale and to start from different venture points, at least for important issues, and when time allows. The very same strategies guard against inappropriate closure and satisficing.

If inputs and expectations fit perfectly, the process comes to a momentary partly standstill and then continues with new input data. If discrepancies surface, they have a strong impact on the elicited actions [5]. In case of severe mismatch, the first schema is altogether discarded and another, new, conceptual frame is tried. The actual appropriateness of a schema can vary over a wide range. In any case, consumption analysis delivers a gradual measure for the goodness of fit between expectations derived from experience and actual inputs.

In addition to the immediate feedback on the quality of fit between expectations and available data provided by the consumption analysis directly, (meta-) information can be obtained by monitoring the general flow of action through the cycle. Such self-monitoring is employed for self-steering according to the Ouroboros Model; the flow of activity and the outcome of the different processing stages exert sequential mutual modulations[3]. In this iterative process, the basis for subsequent steps and action at later occasions are laid.

As to the neural implementation of this central comparison-function, one area in the human brain is the obvious candidate substrate: anterior cingulate cortex, ACC. An influential proposal sees “conflict monitoring” as the main task performed there [26]. At the same time, there is also ample evidence for
response selection [27]. Taken together, the experimental findings can be understood as different facets and perfectly fitting with one consistency checking process of consumption analysis. Whereas a detailed account goes beyond the scope of this paper, for a comparison of expectations with available support in first approximation it suffices to subtract the respective activations in an aligned superposition, which directly highlights their differences, i.e., unassigned and discordant components. Some type of system non-linearity further enhances the contrast, effectively implementing a global winner-take-all behavior, i.e., attention is directed towards discrepancies and, on a more global level, a selected (part of a) schema blocks possible alternatives (Fig. 1). Practice leads to dramatic changes in processing speed, accuracy and required effort, and is accompanied by reductions of activation, similar in cingulate areas and the cerebellum [28]. This is in the focus of work in progress.

3. Concept Formation

Two special types of occasions are specifically marked in the Ouroboros Model as immediately interesting by the outcome of the consumption analysis process, and attention is triggered, which leads to higher than baseline excitement and to stronger activations; preferentially for these cases, new entries are laid down quickly in (episodic) memory[4]:

- Events, when everything fits perfectly; i.e., associated neural representations are stored as kind of “snapshots” of all concurrent activity, making them available for guidance in the future as they have proved useful once;
- Constellations, which led to an impasse or problem, are worthwhile remembering, too; in this case, for future timely avoidance;
- Associations and categorizations are gradually distilled from the statistics of co-occurrences.

Novel categories and concepts can also be assembled on the spot by combining (parts of) existing memory entries following an external trigger[5]. As probably successful and useful repetitions will be pursued more often than futile ones, the third option can be seen as rather similar to the first. Basically these effects, how schemata and novelty and also repeated co-activation can lead to effective storage in long-term memory have been reported as the results of experiments, and they have been taken as basis for a similar model with respect to memory formation [29].

The Ouroboros Model thus can explain how in an evolutionary and self-promoting manner the very memory structures are established such that they naturally serve as the suitable basis for the efficient working of the involved structure-generating processes.

In all cases, the described form of pattern completion enables an agent to advantageously act in anticipation: a few features trigger relevant recollections, and this allows steering subsequent activity to better survive and reach whatever goals. With the improved integration of schemata, a smaller fraction of the components can already trigger the whole gestalt and optimized action can begin faster. As also more and richer schemata will be accrued with experience over time, i.e., concepts with vaster domains and, at the same time, built from more and finer distinguished features, there is a (delicate) balance, which prevents unjustified generalizations from too firmly entrenched action dispositions. The quality of schemata and their optimum interplay in the effective and systematic execution of the processing steps, in particular, self-referential consumption analysis, determine what can be done or thought of efficiently [30].

Coming back to error-handling and repair: no matter how good the respective strategies and procedures are, better it is to avoid mistakes to start with; this is exactly, what according to the Ouroboros Model expectations and anticipations, at least in areas, where an agent knows a little, offer.

Just the same as old memories, new concepts are laid down in the form of cohesive packages, immediately
effective again as schemata, frames or scripts. Building blocks include the total of whatever representations are active at the time when such a snapshot is taken, including sensory signals, abstractions, previously laid down concepts encompassing features relating to probable transients and causal structure, and also prevalent emotions and longer lasting moods. They might, in some cases, but need not, correspond to direct representation-units like words. At subsequent occasions, they serve for controlling behavior, by guiding attention and action towards or away from the marked tracks, depending on the sign of the associated emotion value (which is originally itself distilled from consumption analysis)[5].

4. Need for Interpolation

Some structure and parcellation of all memory content into well separated schemata most probably is dictated quite generally by the need to keep the total amount of stored data manageable. This holds true for quite different levels of abstraction, starting from percepts and extending to abstract theoretical concepts of the highest accessible complexity. Effective categorization is also a consequence of the generation of many entries as snapshots and ad-hoc assemblies as seen from the Ouroboros Model’s perspective. Even in cases when a schematic relation, e.g., representing a movement, is distilled from repeated similar activations, it is most probably laid down economically as distinct (end)points, and the complete transient in between is not stored in arbitrary detail.

While perfectly suited for a process like consumption analysis, a coarse-grained structure of memories poses a challenge when details finer than available in the form of actual local recordings are needed, and in particular, when smooth transients in time are demanded for whatever actions.

Especially with a focus on time, it seems obvious that interpolation can significantly enhance cortical representation capabilities over what is possible with only unitary activity.

4.1 Various Time Scales

The transition between distinct and separated stepping stones, decisive for the overall coherence of activity according to the Ouroboros Model, is affected by diverse processes at different timescales and levels of detail.

Starting from extended timescales, personal features, emotions and moods ensure some coherence and continuity of perceptions as well as for the actions of an agent [5,6].

Over short to medium durations in the order of minutes or seconds, the flow of action according to the Ouroboros Model is mediated by shared constituents, i.e., common attributes and features, of thus concatenated, otherwise distinct and schemata.

Closer to the short-term limit of action, both for bodily movements and also for more abstract cognitive processes, it is hypothesized that representations pertaining to intermediate values are calculated from more directly accessible neighboring reference points by means of some type of averaging and interpolation.

Remarkable, at the other end of the timescale again, given the intrinsic dynamic characteristics of neural action, generating some truly constant level is not completely trivial and appears to mandatorily require some form of integration and averaging.

All timescales are addressed in the Ouroboros Model by schemata including explicitly dynamic features, which code for changes and transients. Their effective resolution will be enhanced by interpolation in turn.

4.2 The Interpolator Hypothesis of the Cerebellum

The cerebellum provides fine-grained values for features in between well established, separately and distinctly represented reference points, i.e., interpolations between cerebral activations, which specify directly available values in a coarser or fragmentary manner.

The cerebellum is thus seen as a dedicated co-processor working in close interplay with the cerebral cortex, greatly expanding the total effective
achievable resolution of representations in living brains.

5. Implementation in Vertebrate Brains

As a theory of human cognition, the Ouroboros Model at some point needs to demonstrate the correspondence of the proposed structures and processes with actual facts as observed in real brains. This is work in progress; first proposals have been presented, e.g., by identifying the hippocampal structures as providing an efficient and rapidly established index to extensive and detailed memory entries, laid down in cerebral cortex[8]. Strengthening the general case for the Ouroboros Model, recent and ongoing experiments yield results concerning the impact of ad-hoc assembled schemata in naive subjects when observing figurative or abstract works of art fully in accord with the predictions [31].

5.1 Selected Hints from Anatomy and Established Findings

First, looking only at a very coarse level, the cerebellum grew in tight lockstep with the cerebral cortex in its evolutionary trajectory in mammals, probably generally in vertebrates[32,33]. Over eons, the cerebellum did not change in its comparatively simple and very regular internal cytoarchitectonic structure in rather diverse vertebrate animals after it once had acquired additional functionality compared to cerebellum-like structures in cartilaginous fish [34]. As the most prominent distinction, cerebellum-like structures do not have climbing fibers. They are clearly sensory, and they act as adaptive filters, which cancel self-generated noise in electrosensory and lateral line systems.

While in humans the majority of the cerebellum is functionally coupled to cerebral association areas, primary visual and auditory cortices, which are largely concerned with feature extraction, appear to be not represented in the cerebellum [35].

A very uniform layout lead early to the suspicion that the cerebellum does one and the same operation to all input arriving there. Some gross correspondence between different areas in cortex and sectors of the cerebellum is observed but the detailed organization into areas is well distinct from what is found in cerebral cortex[36,37]. Input routes and the output tracts establish links in separated closed loops between delimited patches of cerebellar cortex and distinct areas of cerebral cortex; in particular, parts of the body are not represented continuously over an extended area of the cerebellar cortex. Instead, representations are fractured into small discontinuous patches in an apparently uncorrelated manner with diverse sensory and motor areas arranged in close neighborhood[36,38].

Signals carried by relatively small numbers of input and also output fibers are in between expanded enormously with GCs being by far the most numerous neurons in the brain. They are the origin of very many parallel fibers. It has been argued earlier that this stark contrast in numbers allows for (internal) very fine-grained encoding and pattern separation[39]. There is considerable evidence that mossy fiber input codes are preserved in synaptic responses of GCs; this “similar coding principle” is claimed to work as an ideal noise-reducing filter allowing the transmission of weak sensory inputs in a graded fashion[40].

A second input path to the cerebellum runs via the IO (inferior olive), which has been assigned diverse roles in timing by different models; the IO is the sole origin of climbing fibers [41].

Mossy fiber and climbing fiber input pathways from one and the same source, e.g., a point on the skin of an animal, have been shown to converge on the level of single Purkinje cells[42].

Purkinje cells thus receive input via two principal ways, either by a vast number of inputs from parallel fibers piercing their extended dendritic trees (eliciting simple spikes), or, by a single climbing fiber (sparking complex spikes) (Fig. 2). The response of synapses on Purkinje cells to input from parallel fibers is reduced if
this parallel fiber activation “predicts” climbing fiber activity for this cell, i.e., if the latter arrives 50-200 ms later, which actually appears to be a rather long and unspecific interval if one believes in exact timing being the main function of this circuitry [43]. Persistent long term depression of the involved connections ensues after repeated pairings [44].

As shown in Fig. 2, excitatory synapses are denoted by (+) and inhibitory synapses by (-), input involving gap junctions is denoted by arrows. MF: mossy fibers; DCN: deep cerebellar nuclei; CF: climbing fiber; CFC: climbing fiber collateral; GC: granule cell; PF: parallel fiber; PC: Purkinje cell; GgC: golgi cell; SC: stellate cell; BC: basket cell. Effective constant fraction detection is implemented by the dual action of the IO on the DCN plotted as overlay with curved arrows in light blue and green; figure adapted from Wikipedia.

5.2 Effecting Interpolation

Interpolation in the cerebellum is hypothesized as being performed between reference points, i.e., between somehow distinct representations in the cerebral cortex. As a prototypical example benefitting from interpolation, one can take a reaching movement including a start and an end point as a spatial goal (Fig. 3).

It is hypothesized that interpolation is performed in a feed forward manner by determining a shortest trajectory between supporting points in a high-dimensional space defined by the activated feature representations in cerebral cortex and also involving direct sensory input signals. The interpolated values are then relayed back to the (same) concerned cortex areas.

The simplest case would be the determination of a representation for some finely distinguished nuances between the endpoints of a scale for a single (dynamic) variable. When many feature dimensions are considered, derived intermediate values lie in the corresponding hyper plane.

According to the interpolator hypothesis, in the example of a straight movement, all effected activations for starting point and end point, required muscles, expected sensory feedback as well as usual
duration are taken into account; dimensions specifying an abstract goal for actually performing the movement come in addition. All implicated feature dimensions are contributing with some allocated and adjustable weight.

In any case, the resulting trajectory in this high dimensional space will be derived according to a suitably implemented principle of least action. That such computations can in principle be performed by brain circuits have been shown repeatedly; Karl Friston[45] has proposed somewhat related ideas under the name of free-energy minimization.

There exist uniquely outstanding points, i.e., when a given (multidimensional) reference-value and the action-result based on the associated interpolation are exactly equal; the obvious case would be a correctly reached intermediate or end point of a trajectory.

6. Relations to Prevalent Conceptualizations

While not directly following any of the numerous existing proposals concerning the computational functions of the cerebellum (to the best knowledge of the author), it is claimed that the above advanced “interpolator hypothesis” fits rather well with observations and can advantageously incorporate some ideas of all the prevailing models of cerebellar function [39,44,46].

Very sketchy still, a core proposal of the interpolator hypothesis is that the vast number of GCs effectively leads to an only very smoothly changing excitatory input to the still numerous contacted PCs, each influenced by a huge number of very finely graded and reliably coded dimensions. Based on this input, PCs can deliver sustained high action-potential firing rates, which are probably effective as rather constant values in the further processing. PCs provide the calculated interpolated values via DCN to the same cortex areas, which first prompted the interpolation operation.

Special action, as expected for the exact coincidence between any preset value and an interpolation result, in turn, is postulated to be signaled by CF input to PCs. This triggers the PC (complex spike) and shuts it off for...
The Ouroboros Model and the Cerebellum, the “Interpolator Hypothesis”

some time interval immediately afterwards when no addition or correction is required. The timing of the process fits nicely if suitable anticipated values were effective at the outset.

CF activation would thus be foremost a confirmation signal rather than an error signal. As has been pointed out before, “error” and “learning” do not make immediate sense if connections between identical features are affected [42]. In case of strong deviations from expectation, of course, interpolation needs to be stopped, too, and reset [47]. This might explain why CF excitation has been observed in such conditions and identified with a primary error signal.

Learning is held to be different in different phases; for system “set-up” during maturation, or, after a massive change, pronounced long term adaptations can be expected; multiple synergetic processes are thought to be involved. In contrast, during “normal operations” no large persistent modifications in the cerebellum are required or helpful. This matches with the experimental fact that genetically engineered mice exhibit normal motor learning in the absence of long term depression, LTD, at the parallel fiber-Purkinje cell synapse [48]. Consequently, learning, which takes place involving cerebellar activation, might mainly be implemented in the connected cerebral cortex areas [49,50]. To this it also fits that the effects of practice on the functional anatomy of task performance include amongst others a shift of activity away from cerebellum to cortex in rather different tasks [51,28].

For body movements, also when only imagined, distinct postures are separated by a time interval, dictated by basic physics[52]; it is, therefore, clear that a failure to follow a smooth path between them would appear and can be interpreted as a timing problem. Although a definite consensus has not yet been reached, timing has been postulated as a main function of the cerebellum[53-55]. The here presented view of timing errors as only a sequel of disturbed interpolation and resultant faulty coordination does not preclude that proper representations for time intervals appear to be also among the features, which are similarly as any other elaborated (i.e., interpolated) by the cerebellum [56,57].

Serial processing has been postulated to underlie human movement production with temporal segmentation and a series of postures under intermittent control [58-60]. Essential tremor and intention tremor have been linked to cerebellar dysfunction [61]. Interpolation, in any case, will smoothen all effected activities.

Disturbances in postural tone and smooth movement were historically among the first deficiencies associated with cerebellar dysfunction [62]. Complex movements have been described as being broken down into components.

Fig.3 above depicts a comparison between healthy control subjects (traces (a) and (c)) and patients with cerebellar damage (traces (b) and (d)). When asked to move a finger from one point in space (A) to a prescribed goal (B), a healthy subject draws a straight line, whereas patients with a lesioned cerebellum produce wiggly trajectories (A’ to B’), bearing witness to struggling for control and fine tuning. Comparing traces (c) and (d), latency after a go-signal in one and the same patient is higher for the impaired limb, in which the onset of action is reported as coincident with the reaching of the first hold point for the faultless movement. Here, especially the second hint is interesting; this illustration, based on Claude Ghez[62,63], can be interpreted as an action with the impaired limb only starting at the instant when consumption analysis detects a discrepancy, i.e., a deviation from a set goal or reference. The figure is based on the work of Gordon Holmes in the 1920s, and it would be very interesting to scrutinize these old findings[62,63].

Enhanced cerebellar activity is most probably linked to conditions in which wider and more daring interpolations are needed and thus there is a higher risk to make an error. According to the Ouroboros Model, for any deviation between feedback and expectation
consumption analysis triggers attention and also delivers some affective signal (based on computations in the cingulate cortex areas). If the cerebellum is involved in calculating required intermediate values it is no wonder that it is also activated at times when deviations are detected.

Interpolation can also be understood as one facet of predictive coding with activity resulting from a feed forward model [45,64]. With prediction, here, no anticipation of an unknown future is meant but rather the outlining and following of a track in well-charted territory between established reference points. Recently, it has been shown that perception is disturbed in cerebellar patients compared to healthy subjects when proprioception during active movement is demanded, and the performance of healthy controls falls to the level of patients as soon as unpredictable disturbances are applied[65].

Exact timing is persistently being attributed to the cerebellum as one of its main functions. In this respect, it is very enlightening that a part of the detailed circuitry of the cerebellum input tracts can be seen as implementation of a constant fraction detector as used in experimental particle physics for obtaining precise and reproducible trigger signals even in the case of a variable input signal strength [66] (Fig. 2).

The IO contacts the DCN via two routes: directly exciting and with a short delay and inversion via CFs and PCs (Fig. 2). This circuitry looks like a classic example of a constant fraction detector, which triggers DCN output at exactly reproducible times even when there are strong variations in the strength of the arriving signal. Suitable values for the delay duration must be short compared the rise-time or pulse-width of the interesting signal, and they depend on the employed amplification in the enhanced path. CF activation provides the most powerful triggering observed in the vertebrate brain, which dovetails nicely with the demand for fast action under the general conditions of neural circuits.

The very high intrinsic specificity of the cerebellar circuitry can be efficiently modulated [43,67]. Bidirectional, reversible, transient and fast electrotonic coupling between IO neurons via gap junctions in glomeruli effectively widens or shrinks the receptive window for synchronized climbing fiber excitation. It is hypothesized that the ensuing relaxing of coincidence conditions works in accord with greater baseline activity during speed-accuracy tradeoff; both mechanisms in a complementary manner result in faster but less accurate action [68].

The other input path to the cerebellum arriving via mossy fibers from precerebellar nuclei has been discussed for decades, too; it has been shown that under realistic conditions parallel fibers do not generate "beams" of activated PCs but are modulating and balanced by inhibition from interneurons [47] (Fig. 2). In this branch, powerful control modulates summation of parallel fiber input to PCs, proposed to result in effective high pass filtering and allowing spike bursts produced by GCs to preferentially evoke PC single spike output [69]. Filter properties of SCs have been claimed to enhance the contrast of Purkinje cell responses to sparse as compared to clustered synaptic activation from parallel fibers and GCs [70].

Without pretending to already fully understand the exact way in which interpolation is actually performed it can be stated, that the above listed tesserae intuitively seem to fit with a delicately balanced and modulated integration of widely dispersed inputs and the effect relayed back to the connected cortical areas.

Quite generally, failures to provide fitting interpolations would certainly provoke errors. In the realm of abstract cognition, the corresponding errors have been claimed to manifest as "dysmetria of thought" [71]. Cerebellar lesions produce amongst others selective deficits in verbal working memory [72]. In cerebellar patients, working memory, executive tasks and verbal fluency are impaired, and some presumed compensatory recruitment of subcortical brain structures and remote cerebral cortical regions has been observed [73]. Involvement of the cerebellum in
transient and sustained responses in long-term as well as prospective memory has also been reported [74,75]. Following practice, direct cerebral associations and schemata, which are better tailored to the task in question, will diminish the need for interpolation in skilled performance [28,51].

Two special cases can shed more light on the function of the cerebellum.

6.1 Coding of a Constant Value

It has been found repeatedly that during the retention interval together with cerebral activation as well as the cerebellum is active in short term memory tasks[73]. Inclusion of a delay period before recall is especially detrimental to cerebellar patients [72]. The cerebellum is concerned with transients and this can help also with arbitrary successions. Maybe more important, the persistent coding of constant values, i.e., of singular items and the arbitrary combinations of features, e.g., for a mental rehearsal loop to come back to, is very likely supported by tonic cerebellar activations. The latter can naturally be conceptualized as “interpolation” between identical (recurring) reference points. Independent support for the involvement of the cerebellum in working memory comes from the observation that the same gene set, which links neuronal excitability to psychiatric disease, implicates parietal cortex together with the cerebellum [76].

6.2 Interpolation When Input Is Incomplete, Extrapolation

More often than not, reference values for the hypothesized interpolation in the cerebellum will not be completely specified by the actually available neural signals. The cerebellum has been hypothesized to predict time in perceptual events [77]. Here, it is claimed that any dimension can be predicted, i.e., extrapolated, employing the same basic mechanisms as for interpolation. The only difference lies in the availability and completeness of suitable reference values. Depending on the perspective, interpolation could thus just the same be seen as extrapolation in cases with rather complete (forward) models.

Both, interpolation proper and extrapolation, are naturally performed by us in a linear manner, e.g., problems with thermal radiation governed by a law involving the forth power of temperatures can fool the intuition of experts after decades of working in this field[78].

7. Conclusions and Future Work

In a self-reflective and self-relational consistent way the Ouroboros Model holds that cognition generally progresses in an interplay of top-down “frame-setting” and bottom-up “filling-in of slots”. Accordingly, emphasis lies on first sketching an overall conceptual frame of reference before digging into intricate subtleties. It is claimed that conceptual work and progress is a mandatory prerequisite for successful and efficient detailed investigation and modelling to follow thereafter. A case in point here is the diversity of extant models of cerebellar function [39,44,46,54]. Formalized and rather specific, they allow modeling of selected findings and yet, these best that we have to date make some diametrically opposed (quantitative) predictions leaving us in the unhappy situation that we obviously do not yet fully understand what the numerical majority of the neurons in a human brain are good for.

No doubt, concrete and quantitative formulations and also working implementations are required at some point in time, and extensive detailed work is still needed. In particular, formalization and numerical simulations are necessary to illuminate details concerning timing and how in a biologically plausible manner interpolation is actually implemented, most probably following a principle of least action, with the neural networks found in the cerebellum. Collaborations to formalize the Ouroboros Model in general and the interpolation function, in particular, and to tackle these issues by simulations would be most welcome.
As one example, for functional activation studies, predictions can be made concerning differences and similarities of cerebellar contributions to movements, e.g., comparing the drawing of a complex figure either by hand and with a pencil or with a big brush and using wide arm movements: timing and also the involvement of the body would differ vastly but in the proper reference frame the abstract specifications for supporting points and their smooth interpolation, i.e., transients, should be rather similar.

Pertaining to abstract cognition, it can be predicted that cerebellar patients are particularly deficient when asked for estimates, extrapolations beyond known details and/or interpolations between known (and little related) facts.

The interpolator hypothesis presents a new, overarching and presently coarse-grained picture. Preliminary evidence is presented here that the cerebellum serves useful and deemed necessary functions as an interpolator for deriving fine-grained representations and smooth transitions from distinct supporting points defined by activations in cerebral cortex and effectively referring to different points in time.

Although the idea and motivation comes from a specific cognitive architecture, the “Ouroboros Model”, the proposal is not really dependent on this basis. The outlined account can explain the benefits of cerebellum-like structures and of a full clowm cerebellum in primitive animals for movement and sensing, the observed reciprocity between, and apparently parallel expansion of, cerebellar and cerebral cortex, and, in particular, the generally observed link to timing, the latter not only pertaining to motion but also to perception and even abstract cognition. It is sketched how an all-embracing conceptualization, i.e., the interpolator hypothesis, not only can explain available observations but also reconcile several diverse approaches and distinct earlier proposals documented in the literature. It is anticipated that the interpolator hypothesis leads to a better understanding of healthy and impeded human cognition and thus strengthening the basis for addressing the BICA challenge [2]. A centralized dedicated interpolator function might be widely useful in many diverse cognitive architectures in a similar vein as has been claimed for meta-cognition and self-monitoring [6,24,30].

A partial implementation of the Ouroboros Model in a safety installation is working in principle; interpolation has not yet been fully included due to unrelated problems; continued work is in progress [79].

Fully in accord with the spirit of the Ouroboros Model, one can also make the point that the interpolator hypothesis is an example of “turning the wheel another round”, i.e., a biological inspired model of cognition draws attention to specific questions and subsequently leads to testable predictions concerning the function of the numerical majority of the neurons in a human brain.

Acknowledgements

A number of insightful comments, hints and very helpful suggestions for improvements by several colleagues are gratefully acknowledged.

References

[7] K. Thomsen, Is quantum mechanics needed to explain

The Ouroboros Model and the Cerebellum, the “Interpolator Hypothesis”
The Ouroboros Model and the Cerebellum, the “Interpolator Hypothesis”


The Ouroboros Model and the Cerebellum, the “Interpolator Hypothesis”


