A Theory of Cerebral Information Processing

Sigurd Enghoff
enghoff@scientific-concepts.com

Abstract - **The author proposes an information processing framework, which is theorised to underpin information processing in the cerebrum. The framework focuses on self-organising mechanisms in the cortex and on interactions between cortical and sub-cortical processes that give rise to sensory perception, motor control, memory, learning, planning and problem solving in the cerebrum. Salience and priming are argued to be core mechanisms of an attention system, which governs the routing of information among cortical areas and is central to memory consolidation.**
It is argued that as a consequence of the framework, explicit and implicit memories share common neural mechanisms and that their apparent distinct characteristics result from properties of awareness rather than differences in neural circuitry.
In the concluding section, the framework is considered from an engineering perspective. An information processing model is proposed, which divides the framework into functional components, which are rooted in statistics, robotics, artificial intelligence and control theory.

Any feedback in opposition or support of the presented theory would be much appreciated. Kindly send any correspondence to the author to Sigurd Enghoff at enghoff@scientific-concepts.com

Introduction

The cortex is known to feature significant anatomic uniformity (Rockel et al., 1980). Barring a few exceptions, neurons are arranged in 6 vertical layers from the surface of the cortex and exhibit a columnar organisation that repeats across the horizontal plane of the cortex (Kolb et al. 2003). It is expected that the anatomical uniformity underpins equally uniform functional and processing mechanisms across the cortex. Thus our understanding of the information processing performed by cortical regions as the prefrontal cortex and cortical motor areas may be improved by relating to our knowledge of the processing performed by the cortex in other areas as the visual cortex.
The present publication takes an altogether novel approach to neural information processing and considers mechanisms of information processing in the cerebrum from a functional perspective.

Sensory Processing

Arguably the primary visual cortex and the extra-striate visual cortical areas are among the most extensively studied and best documented regions of the cortex. Studies of the visual cortex provide good evidence that the cortex is highly effective at extracting and separating features present in a visual scene, with higher visual areas representing increasingly more abstract features and locational information about visually observed objects.

Signal separation and classification are central to processing performed by the visual cortex and other sensory cortical areas. Mini columns encompassed within a hyper column share common innervating input (Mountcastle, 1997; Buxhoeveden et al., 2002). Through a mechanism of mutual competition, mini columns within a hyper column self-organise such that each mini column responds to a unique feature in the input.

In addition to the forward pathways, which separate and classify signals as they propagate through the hierarchy of sensory areas, the cortex has extensive reciprocal connections from higher to lower cortical areas. Also extensive connections exist between different sensory modalities and across cortical areas, which makes the cortex a highly interconnected network. Relationships between information in different cortical areas are established as part of the memory consolidation process described in later sections. The ability of the cortex to establish diverse and strong relationships between information across diverse cortical areas, gives the cortex a highly effective means of basic as well as abstract association. E.g., the voice of a relative may be associated with the visual representation of the person's face and with the emotional relationship that is had with the person.

Motor Control

As stated in the introduction, it is anticipated that the fundamental processing done by the cortex is largely uniform across cortical areas. Thus the signal separation, classification and pattern association performed by the cortex in sensory areas would also underpin information processing performed by the motor areas.

While cortical columns in sensory areas represent features of observed stimuli, cortical columns in motor areas represent action schemas. Sensory features are learned through observation of environmental stimuli, correspondingly action schemas are learned by observing and capturing correlations between initiating conditions for an action and conditions causally resulting from the action.
E.g., a schema for bowling a ball will associate proprioceptive and spatial information about the configuration and location of the body and ball prior to the execution of the bowl, with the anticipated state of the body and ball subsequent to the bowl. Changes in body posture resulting from the bowl are captured by the cortex and consolidated along with an association to a set of motor schemas, which comprise the bowling action. For a relatively complex schema as a bowl, connections with parietal areas will capture spatial translations as the trajectory of the ball. Abstract schemas may form highly diverse associations across the cortex in order to capture the anticipated effects of their execution. Action schemas are organised in a hierarchical fashion with increasingly more complex and abstract actions at higher motor areas.

Relationships between the prerequisites for an action schema and its corresponding consequences are learned by capturing and storing observed causal relations and correlations. Learning and consolidation of action schemas in the motor cortex happen through the same mechanisms that drives learning in other cortical areas. (See section on Memory Consolidation)

The motor cortex is very capable at defining targets, planning actions, monitoring action execution against targets and dealing with contingencies; however cortical processing is elaborate and relatively slow, which makes the cortex inept at rapid and precise motor control that requires tight closed-loop regulation. For actions that call for accurate timing or highly responsive dynamics, the cortex is expected to engage subcortical circuitry and in particular the cerebellum. A central role of cerebellar circuits is to capture and model kinematics of the skeletomuscular system, which places the cerebellum ideally to exhibit tight closed-loop control over the execution of motor actions.

Memory Consolidation

In the subsequent sections the term *percept* will be used to describe complex and possibly multimodal stimuli, which correspond to single coherent observations.

Although subcortical areas, and in particular the hippocampus, guide learning and memory consolidation in the cortex by directing information flows, learning in the cortex at the level of cortical columns is expected to occur without much if any direct supervision. Consequently the cortex does not receive signals to explicitly trigger learning or reinforcement, rather the cortex self-organises around the signal patterns, which are observed locally.

A central function of the hippocampus is the capture, storage and recall of salient percepts (see section on Salience) that arise in the cortex while an organism is awake. The hippocampus has the ability to reciprocally innervate the cortex and re-establish captured percepts, thus yielding the recall of previously stored percepts. The hippocampus employs this recall mechanism during sleep, which enables percepts that were captured and stored in short-term hippocampal correlational patterns to be reliably transferred and encoded as long-term memories in the cortex. For the long-term encoding of memories, the hippocampus relies on the intrinsic ability of the cortex to self-organise and adapt to patterns and percepts that the cortex encounters.

As percepts are encountered and accessed repeatedly over longer periods (days or weeks), the self-organising adaption mechanisms gradually dedicate neural circuitry to the representation of common percepts. This process causes frequently encountered percepts to require progressively simpler and sparser cortical patterns for their unambiguous representation, while increasingly dedicated cortical circuitry becomes devoted to the representation of learned percepts. The self-organising process that drives the incorporation of new percepts and patterns builds upon existing cortical pattern decompositions. As a result the adaption process intrinsically structures pattern decompositions into hierarchically organised assemblies.

Salience, Attention and Learning

At the columnar level the cortex is expected to be capable of capturing and learning coherence patterns. Individual cortical columns adapt to associate with a unique pattern, which may be a feature in a sensory area, schema in a motor area, rule in a higher planning area, etc. The cortex is anticipated to acquire new patterns and adapt to new information through self-organisation, which at the columnar level is driven by competition among columns.

While adaption and learning at the level of cortical columns is unsupervised, routing and prioritisation of information flows among cortical areas are expected to be orchestrated by centralised mechanisms. Since columns self-organise around signal patterns that are received locally, they are exclusively dependent on the information that they are fed. Consequently the central information routing mechanisms are able to guide learning in the cortex by controlling signal flows.

The policies that govern prioritisation and routing of information at any given instance are dependent on the mental process being executed. Policies associated with attention, as salience and priming, guide information capture, processing and storage and these policies ensure cortical resources attend coherently to information of the highest priority. Other policies guide the routing of information for problem solving and planning, while yet different policies direct information flows during sleep. The policies are implemented by a multitude of sub-cortical networks while the execution of the policies and the routing of intra-cortical signals is accomplished by the thalamus.

Salience

Cortical hyper columns are expected to generate signals that reflect the degree of salience that a column observes in its input. When a hyper column receives a pattern that it is able to readily anticipate and confidently classify the corresponding salience value is low, while a pattern that is unexpected or appears ambiguous will produce a comparatively high salience value. An alternative interpretation of the salience value is as an estimate of how well the cortical "model" fits the presently observed input data, which in technical terms makes salience a measure of the posterior entropy contained in a pattern provided the cortex as a model. In sensory areas, salience reflects the level to which observed sensory stimuli match known and anticipated features, while in planning and motor areas salience represents the degree to which learned rules and causal relationship agree with environmental observations.

Patterns that produce high salience values are prioritised over less salient patterns as they are communicated between cortical areas. Thus highly salient patterns are likely to form percepts that in turn impact prioritisation of planning and actions. Salient patterns are also captured and encoded by the hippocampus, which allows the captured patterns to be subsequently recalled and reproduced. Pattern recalls are essential for planning, problem solving and sleep processes. (See section on Memory Consolidation)

Priming

Priming is an anticipatory tuning of cortical columns to toward particular patterns. While salience is a function of information that is encountered and observed, priming of the cortex occurs in advance of an expected stimulus and makes the primed cortical columns particularly sensitive to the priming patterns. When a primed cortical area receives a pattern that matches the priming pattern, the cortex will attribute a high a salience value to pattern as it is processed and conveyed. Conversely, when a primed area receives patterns that differ from the priming pattern, a low salience value will be attributed and the pattern may be blocked from further cortical processing.

Priming patterns are communicated between cortical areas and will typically cascade from higher to lower cortical areas. Priming is governed by policies and heuristics, which are enforced by sub-cortical networks as it is generally the case for intra-cortical communication.

Salience and priming affect cortical information flows and pattern processing at the level of cortical columns, at relatively short time scales (hundreds of milliseconds) and are principally concerned with "reactive" processing of information. Thus salience and priming have no long-term effects on their own, although they are central to other processes that do have lasting impact. Salience and priming at cortex, in conjunction with their governing sub-cortical processes, are core elements of attention.

Classes of Memory

While different classes of memory, as declarative, procedural and spatial memories, are all acquired through the capture of percepts, each class of memory has its primary root in a distinct cortical area. I.e., declarative memories are centred in sensory visual and verbal areas, procedural memories in the motor areas, spatial memories in parietal areas, etc.

The proposition that all cortical memory is captured and stored through the same principal mechanism rests on an assumption that explicit and implicit memories are quantitatively identical. It is argued that the distinction between explicit and implicit memories arises because we are more explicitly aware of percepts present in certain cortical areas than others. Specifically, we are most acutely and explicitly aware of percepts that relate closely to observable aspects of our environment, such percepts typically engage higher sensory areas in the cortex, particularly higher visual and verbal areas. Memories that are deemed explicit have representations that evoke these same cortical areas.

Conversely, implicit memories have representations in cortical areas that process comparatively more abstract information and these areas do not form part of awareness.
In summary, it is argued that the distinction between explicit and implicit memories derives from the characteristics of awareness rather than any differences in the neural circuitry or mechanisms that implement each class of memory.

Planning and Problem Solving

The cortex forms associative relationships between information encoded in different cortical areas. As the cortex self-organises around information that is captured and learned, associations are established across sensory modalities and between different cortical areas.

Employing the same mechanisms as sensory cortical areas, the motor cortex captures correlations through passive observation. Specifically, the motor cortex is believed to capture relationships between conditions required for an action and the conditions expected as a result of actions. The captured relationships effectively become action-specific rules, or schemas, as they are consolidated in the cortex.
Conditions are represented by the same principles as percepts in the sensory cortex, thus conditions are stored as relationship between cortical motor areas and areas that are affected when motor schemas are executed. E.g., if a low-level action as the effect of raising an arm, the conditions for the action would involve areas in the corresponding motor sensory cortex. A slightly higher level action schema may relate to raising an arm to catch a ball in flight, the conditions relating to this action would be expected to innervate both motor sensory and parietal areas that are affected by the execution of the schema.
Planning and problem solving build on these diverse associative relationships in the cortex and rely on sub-cortical mechanisms to guide associations in a targeted fashion.

The planning and problem solving process is initiated by sub-cortical heuristics, which prioritise and determine an objective for the process. Higher cortical areas are primed with a target pattern that represents the objective state. An objective may be concrete, as satisfying a state of hunger, or abstract, as solving a jigsaw puzzle. Sub-cortical policies will subsequently drive successive cortical associations toward satisfying the primed target. Planning and problem solving involving action sequences will engage the causality rules of the motor areas, while other cortical areas will be employed depending on the plan or problem at hand, e.g., for solutions to spatial manipulation problems, parietal areas would be engaged.

The association processes, which form the basis of cerebral planning and problem solving, in effect implement the direct search functionality of an automated planner. In fact, employing guided associations to solving planning problems may be a very effective route to solving planning problems in general. Automated planning is discussed in more detail in the subsequent section on the Information Processing Model.

Information Processing Model

The cortex is highly effective at signal and pattern classification as traditionally implemented by statistical techniques developed for cluster analysis and blind signal separation. At the columnar level, the cortex self-organises to separate distinctive and commonly encountered features.
Cortical learning is supported by a salience mechanism that captures and temporarily stores patterns that are deemed salient/high entropy for later recall and permanent encoding. Information in the cortex organises into a hierarchy of increasingly abstract representations. Reciprocal relationships within the structure of the cortical networks, allow information stored in the cortex to be accessed through association. This processing and storage model is expected to be the foundation of cortical information processing across all sensory modalities, schemas for motor control and rules for planning and problem solving.

Hierarchical Task Network

The presented information processing framework proposes that problem solving and action planning in the cerebrum share several key functional and organisational features with Hierarchical Task Networks (Nejati et al., 2009; Sterren 2009; Kivinen et al., 1992). HTNs define individual tasks by a set of pre-conditions, which must be satisfied for the task to be executed, and an associated set of post-conditions, which will become satisfied, when the task completes. The definition of tasks by the pre and post-conditions associated with their execution corresponds closely with the action-specific pre and post-conditions that the cortex is expected to capture and store.
Hierarchical Task Networks, as the term suggests, organise task definitions hierarchically in a structure equivalent to the organisation of action schemas in the motor cortex. Both HTNs and the cortex represent action/task primitives at the bottom of the hierarchy while increasingly more advanced and composited

actions/tasks are represented hierarchically as sets of primitives and schemas.

A HTN defines a goal as a set of conditions that must be fulfilled to achieve a target and an automated planner is used to search the HTN for sequences of tasks, which when executed, will bring the agent to satisfy the goal conditions.
HTNs employ the formalised logic of an automated planner to search the task space for solutions to satisfy the goal condition, whereas the cerebrum relies on a guided association process to determine action sequences that meet a primed target state.

Partially Observed Markov Decision Process
The information processing performed by the cerebrum shares several significant features with a Partially Observed Markov Decision Process (POMDP). POMDPs are frequently used for applications in artificial intelligence and automated planning. The POMDP model deals with the decision making process of an agent in an environment, which is only partially observable by the agent (Theocharous et al., 2001). Thus a central responsibility of the cerebrum corresponds to the function of an agent guided by a POMDP.

The hierarchy of feature representations, schemas and rules that are learned and encoded in the cortex, constitutes a joint environmental and agent model from the perspective of a POMDP agent.

A POMDP driven agent builds and maintains a 'belief state' of its present environment based on a historic trail of observations. The belief state definition is effectively a formalisation of the percept term that has been used extensively in this publication.

An agent governed by a POMDP evaluates action policies by simulating potential belief states, which in turn are used to predict the outcome of future actions. The processes of simulating belief states is qualitatively analogous to the association process, which drives planning and problem solving, as proposed in the previous section.

POMDPs typically call for the definition of a value or reward function that is maximised across available actions to derive an optimal policy for the agent. While the value function may not have a direct cerebellar equivalent, an extrapolation of the value function may assist the formulation of effective control logic required to guide information flows during planning and problem solving associations.

Summary
It is argued that notable parallels exist between information processing in the cerebrum and techniques commonly employed for guidance of autonomous agents with these techniques deployed in close union with statistical methods for signal and data analysis.
A framework is proposed that at its core employs a hierarchical network of statistical clustering processes, which extract and capture correlations in observed data. The network is capable of establishing extensive and diverse relationships across modalities. A generalised POMDP operates by governing and prioritising information flows in the network and the POMDP ensures the maintenance of a present belief state. In a structure akin to a HTN, the network encodes a hierarchy of rules that range from simple motor action to complex schemas and abstract rules. Planning and problem solving is conducted by the POMDP by driving associations in the network that simulate sequences of belief states. Lastly the POMDP monitors execution of plans.

Future Research and Development

Presently a technology demonstrator is in development. The objective is to prove the feasibility of using a single network of self-organising units to both extract prominent features of an observed environment and learn causal relationships between initiating and resulting conditions for motor actions. The present development uses a hierarchical network of Gaussian mixture units. The Gaussian mixture assumption benefits from having very robust statistical and mathematical underpinnings and it provides a good balance between model complexity, flexibility and generality.
As development of the system progresses and increasing levels of hierarchical layers are incorporated, the policies governing information routing between units in the system are expected to require considerable refinement.
From a cerebral point of view, a significantly more detailed theory of subcortical control and information routing mechanisms is required.

References

Atif H. G., Mikko L. H. (2009) Cortical Columns: Building Blocks for Intelligent Systems. *IEEE Symp. on Computational Intelligence for Multimedia Signal and Vision Processing.*

Silvia Chiappa, Jens Kober, Jan Peters. Using Bayesian Dynamical Systems for Motion Template Libraries.

Advances in Neural Information Processing Systems (NIPS) 2009.

Jens Kober, Jan Peters. Learning Motor Primitives for Robotics. *IEEE Intl. Conf. on Robotics and Automation 2009.*

William van der Sterren. Multi-Unit Planning with HTN and A*. *AIGameDev Paris Game AI Conf. 2009.*

Negin Nejati, Pat Langley, Tolga Konik. Learning Hierarchical Task Networks by Observation. *Intl Conf. on Machine Learning 2006.*

E. Todorov, Z. Ghahramani. Unsupervised Learning of Sensory-Motor Primitives. *IEEE Intl. Conf. on Engineering in Medicine and Biology Society 2003.*

Bryan Kolb, Ian Q Whishaw. Fundamentals of human neuropsychology. *New York: Worth 2003*.

Daniel P. Buxhoeveden, Manuel F. Casanova. The minicolumn hypothesis in neuroscience. *Brain, Vol. 125 #5, pp935-951, Oxford University Press, May 2002.*

Georgios Theocharous, Khashayar Rohanimanesh, Sridhar Mahadevan. Learning Hierarchical Partially Observable Markov Decision Process Models for Robot Navigation. *IEEE Intl. Conf. on Robotics and Automation 2001.*

Sethu Vijayakumar, Stefan Schaal. Locally Weighted Projection Regression. *Intl. Conf. on Machine Learning 2000.*

Kevin P. Murphy. A Survey of POMDP Solution Techniques. *2000.*

Vernon B. Mountcastle. The columnar organization of the neocortex, *Brain, Vol. 20 #4, pp701-722, Oxford University Press, April 1997.*

Jyrki Kivinen, Heikki Mannila, Esko Ukkonen. Learning Hierarchical Rule Sets. *Annual Workshop on Computational Learning Theory 1992.*

A. J. Rockel, R. W. Hiorns, T. P. S. Powell. The Basic Uniformity in Structure of the Neocortex. *Brain 103 (2): 221-244 1980*.