Cognitive architecture with visuo-motor and advanced cognitive functions (D3.8 - SGA3)

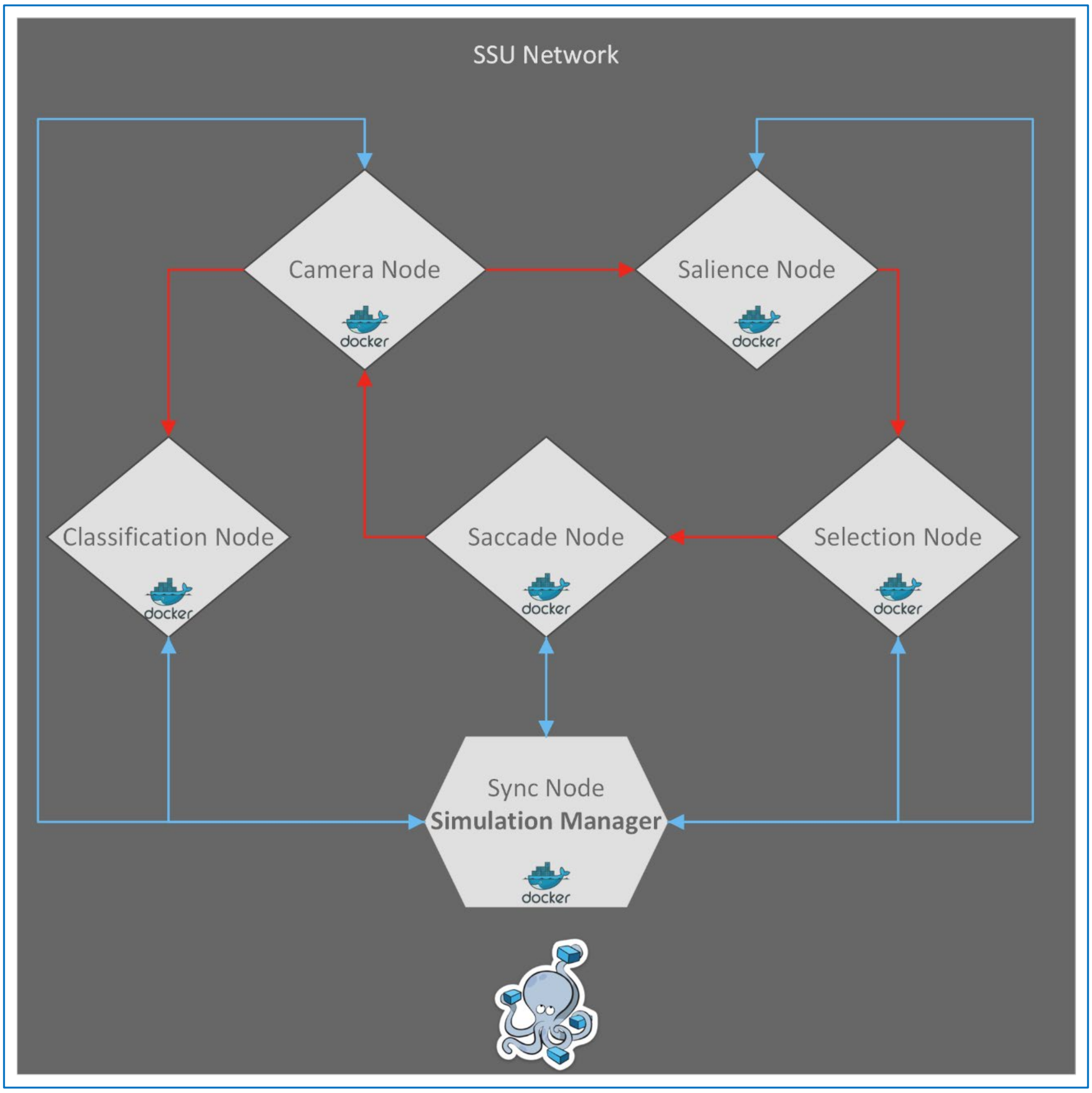

Figure 1: Diagram of the saccades for scene understanding (SSU) architecture.

The SSU consists of five modules (nodes): Camera, Salience, Selection, Saccade, and Classification. An additional Sync Node manages simulation of the SSU. All nodes are isolated within their own Docker container and interact with each other through a virtual network. Containers and the network are orchestrated by Docker Compose. Red arrows show the flow of data through the SSU architecture. Blue arrows show the flow of control from the sync node to the other nodes. Nodes communicate with each other using ROS2 in a publish-subscribe pattern.

Project Number:	945539	Project Title:	HBP SGA3

Document Title:	Cognitive architecture with visuo-motor and advanced cognitive functions
Document Filename:	D3.8 (D27) SGA3 M42 RESUBMITTED 231208.docx
Deliverable Number:	SGA3 D3.8 (D27)
Deliverable Type:	Report
Dissemination Level:	PU = Public
Planned Delivery Date:	SGA3 M42 / 30 SEP 2023
Actual Delivery Date:	SGA3 M42 / 19 SEP 2023 (resubmitted 08 Dec 2023)
Author(s):	Mario SENDEN, UM (P117)
Compiled by:	Mario SENDEN, UM (P117)
Contributor(s):	
WP QC Review:	Anita VAN OERS, UM (P117)
WP Leader / Deputy Leader Sign Off:	Rainer GOEBEL, UM (P117)
T7.4 QC Review:	N/A
Description in GA:	A report detailing the architecture, biologically plausible training procedure, embodiment and performance of a biologically detailed network architecture able to perform visuo-motor and complex A.I. tasks by employing hierarchical cognitive processing. The report will further include ongoing comparisons of brain activity generated by the architecture against neuroimaging and behavioural data.
Abstract:	This report discusses the development of large-scale architectures with visuo-motor and cognitive functionalities through two primary approaches: the modular-integrative approach and end-to-end training via deep reinforcement learning. Our modular-integrative modelling decomposes a complex system into independent, task-specific modules, promoting continuous model refinement and adaptive system evolution. Alongside, we present and end-to-end training approach for developing large-scale of brain models. It provides an extension of goal-driven modelling from sensory tasks to sensorimotor control and cognition, enabling a computationally efficient closure of the perception-action loop.
Keywords:	Large-Scale Architectures, Visuo-Motor Functions, Cognitive Functions, Modular-Integrative Modelling, Deep Reinforcement Learning
Target Users/Readers:	Computational neuroscience community, computer scientists, HPC community, neuroinformaticians, neuroscientists, platform users

Table of Contents

Table of Figures

History of Changes made to this Deliverable (post Submission)

Date	Change Requested / Change Made / Other Action
19 SEP 2023	Deliverable submitted to EC
	Resubmission with specified changes requested in Review Report Main changes requested: • The quality of the text is not sufficient and needs to be revised. • Either add details from or clearly list links to peer-reviewed publications or preprints related to the reported outcomes.
	Revised draft sent by WP to PCO. Main changes made, with indication where each change was made: • Improved writing (all Sections) • Added references of publications related to the reported outcomes. (see Sections 2.3 and 3.3).
	Revised version resubmitted to EC by PCO via SyGMa

1. Introduction

The complexity of the human brain, estimated to contain 86 billion interconnected neurons, poses an immense challenge for neuroscience. Experimental methods alone are limited in their capacity to fully capture and explain the intricate workings of the brain and are increasingly complemented by computational models that describe neural systems in abstract, mathematical terms. These models allow for the simulation, prediction, and exploration of brain function in ways that were previously unimaginable.

However, the development of computational models of the brain is a challenging undertaking. Researchers need to balance the need for capturing real, biological phenomena in the brain and a model's ability to execute complex cognitive tasks. This requires a careful balance between bottom-up and top-down modelling techniques. Hybrid models, that combine these two approaches, show great promise, and enable the integration of existing model implementations into a single system.

To build such hybrid models, we developed a modular-integrative modelling approach that divides a complex system into discrete, self-contained modules, each simulating a circumscribed brain structure or specific aspect of brain function. This not only allows for the independent development and optimization of each module but also facilitates adaptation of the overall system to meet the demands of different research questions. Moreover, with a deeper understanding of the brain's intricacies and the emergence of new hypotheses, the modular approach fosters the examination of these hypotheses as independent modules, promoting constant model improvement and collaborative development.

In addition to the modular-integrative approach, we also highlight the advantages of end-to-end training of large-scale architectures through goal-driven reinforcement learning. To that end, we utilize AngoraPy, a tool developed within Showcase 5 with the specific aim to train brain models capable of simulating a wide array of sensorimotor tasks. Driven by requirements of the present demonstrator, AngoraPy was adapted to provide a platform for goal-driven modelling that not only accommodates sensorimotor control but also extends to cognition. This allows us to close the perception-cognition-action loop in a computationally efficient manner. AngoraPy is user-friendly and has been optimised for distributed computation, scaling from individual workstations to high-performance computing clusters. Furthermore, it adheres to a neuroscience-first approach, prioritising the needs of neuroscientists while also preserving options for customisation. These attributes make AngoraPy an instrumental resource in the exploration of brain functionality and the development of high-performing, goal-driven models.

2. Modular-Integrative Modelling

2.1 Motivation

In the contemporary landscape of neuroscience research and computational modelling, the need for complex, scalable, and highly customisable modelling solutions is rapidly intensifying. Bottom-up (data-driven) modelling provides a biologically plausible but functionally limited perspective, whereas the top-down (hypothesis-driven and goal-driven) modelling offers functional performance at the expense of biological realism. Hybrid models make it possible to blend biological plausibility and functional performance by integrating pre-existing model implementations into a unified system. A flexible, modular-integrative approach additionally supports continuous model refinement and comparative hypothesis testing.

The beauty of this approach lies in the autonomy it bestows upon each module - they can be developed independently, encapsulating specific functionalities or simulating particular biological phenomena. This autonomy offers significant advantages. Firstly, it enables developers to concentrate on specific components, refining and optimising them without the obligation of managing the entire system. Secondly, it allows the addition, removal or substitution of modules to match the requirements of the research question or the data available, substantially enhancing the flexibility and adaptability of the system.

As our understanding of the brain develops and new hypotheses arise, this modularity enables us to construct and assess these hypotheses as separate modules without the need to redesign the whole model. This approach also allows for parallelised development, where different teams can work on different modules at the same time, reducing development time and fostering collaboration.

The relevance of this modular-integrative modelling approach extends far beyond the sphere of academic research and provides a strong framework for commercial use, particularly in the field of artificial intelligence (AI). AI and machine learning are increasingly dominant in various sectors, and there is an unprecedented demand for adaptable, scalable and efficient modelling methods, such as ours.

To increase understanding and promote wider use of this method, we have submitted an opinion piece to *National Science Review* detailing the advantages and applications of the modular-integrative modelling technique. The paper provides finer details of this approach and highlights its significance in current neuroscience research.

2.2 Technical Implementation

Realising the modular-integrative modelling approach involves a multi-step process that includes component modularisation, containerisation, communication via a message broker, time synchronisation with a simulation manager, and synchronisation of components.

The initial step, component modularisation, breaks the system down into manageable modules, each serving a specific function. The succeeding step, containerisation, is achieved using technologies like Docker that isolate each module within a container, circumventing dependency issues and ensuring that modules can function in diverse environments. Efficient inter-module communication is critical and facilitated by a message broker. Lastly, synchronisation is key to ensuring smooth and coordinated functioning of the modules, which we achieve using a dedicated Simulation Manager. Each module signals the Simulation Manager upon completion of its calculations for the current time epoch, and a new epoch is initiated by the Simulation Manager once it receives signals from all modules.

We provide a technical demonstration of the modular-integrative approach for a closed-loop architecture engaging in saccades for scene understanding (SSU; https://github.com/ccnmaastricht/SSU). Scene understanding is a complex task requiring the interplay of numerous modules - each processing a different aspect of visual input, coordinating eye movements, decision-making and more. It thus offers an ideal test bed for a modular-integrative approach. The SSU architecture consists of five modules (see Figure 1). A camera module provides snapshot from panoramic scenes. These scenes were selected from the Stanford 2D-3D-Semantics Dataset (http://buildingparser.stanford.edu/dataset.html) and include 11 categories of indoor scenes. Snapshots of the scene are passed on to a classification module. This module first resamples the snapshot according to ganglion cell distributions in the retina adapted from a procedure outlined by da Costa et al. (2023; P3995; original implementation: https://github.com/ccnmaastricht/ganglion_cell_sampling). Subsequently, it updates the internal state and classification layer of a recurrent convolutional neural network implemented with PyTorch. The camera also passes the snapshot to a saliency module. This module uses a deep encoder-decoder architecture developed by Kroner et al. (2020; P1761; https://github.com/alexanderkroner/saliency) to compute the salience distribution across the snapshot. Then, the module integrates this local, eye-centred, saliency into a global (head-centred) map. Regions of this global map that are not currently in view exhibit passive decay with a time constant of 200ms. The saliency module passes the global saliency map on to the target selection module, which probabilistically selects a target for the next saccade and passes this on to the saccade generation module. The saccade generation module uses a spiking-neuron model of the reticular saccade generator implemented in NEST (https://github.com/ccnmaastricht/spiking_saccade_generator) to update eye-position. The target location and actual eye position are passed to the camera module, which suppresses snapshots while the eyes are in motion. All modules are isolated within a dedicated Docker container and communicate via a shared network. Data exchange between modules is decentralised through a publish-subscribe pattern realised via ROS2. A dedicated sync_node serves as the Simulation Manager to synchronise all simulations.

2.3 Publications

Senden, M., van Albada, S. J., Pezzulo, G., Falotico, E., Hashim, I., Kroner, A., Kurth, A. C., Lanillos, P., Narayanan, V., Pennartz, C., Petrovici, M. A., Steffen, L., Weidler, T., & Goebel, R. (2023). Modular-Integrative Modeling: A New Framework for Building Brain Models that Blend Biological Realism and Functional Performance. Manuscript submitted for publication, National Science Review.

3. End-To-End Training of Large-Scale Architectures Capable of Visuo-Motor and Cognitive Functions

3.1 Motivation

Goal-driven deep learning presents a promising framework for providing a holistic perspective of the brain as an integrated system. Using this approach, cognitive processes, sensory inputs, and motor actions are not required to be separate operations but can be from the outset interwoven components of a single, complex process. The core idea of goal-driven deep learning is that deep neural networks can develop computational strategies that match those employed by the brain. Specifically, if similar perceptual, cognitive, or motor tasks are assigned to these networks as a neural system, they are expected to converge on comparable neurocomputational solutions employed by these biological systems. Historically, this modeling approach has had its restrictions due to its reliance on significant amounts of labeled data, confining its use mainly to areas of perceptual neuroscience where abundant data of this kind is available.

However, reinforcement learning renders the goal-driven deep learning approach viable for developing models of sensorimotor and cognitive functions. Reinforcement learning is distinct in its reliance on learning from data sourced directly from environmental interactions. This approach sidesteps the constraints linked to the necessity for pre-labeled data sets, facilitating a more organic approach to neural modeling. By creating a training framework based on RL techniques, we complement the modular-integrative approach to produce holistic brain models that span the entire perception-cognition-action cycle.

3.2 Technical Implementation

End-to-End training of Large-Scale Architectures is realised with AngoraPy (https://doi.org/10.5281/zenodo.7770180). AngoraPy's prime function is to train deep neural network models of the human brain using reinforcement learning. The design principles that govern AngoraPy—Neuroscience First, Modularity, and Pragmatism—ensure that it remains a flexible yet effective tool for neuroscientists. With a 'neuroscience-first' approach, AngoraPy addresses the needs of neuroscientists aiming to build goal-driven models with ease. Its modularity permits a wide range of applications and caters to various tasks and models. And finally, the principle of pragmatism ensures a balance between computational efficiency, performance, and the simplicity of the API, delivering a tool that is both practical and powerful. For further information on technical implementation of AngoraPy and a demonstration of its capabilities for developing models of sensorimotor integration, see Deliverable D3.14 on Showcase 5.

Here, we demonstrate the ability of an end-to-end goal-driven reinforcement learning framework (realized with AngoraPy) to develop models capable of cognitive function. To that end, we trained a network on the Tower of Hanoi task. For this task, the network architecture consisted of a simple Long Short-Term Memory (LSTM) network with two upstream fully connected layers of 64 units each, leading to downstream policy and value heads. The policy was discrete, reflecting the discrete movements of the disks in the task. The trained model converged to an optimal solution in about 20 training cycles of refinement/exploration (see Figure 2).

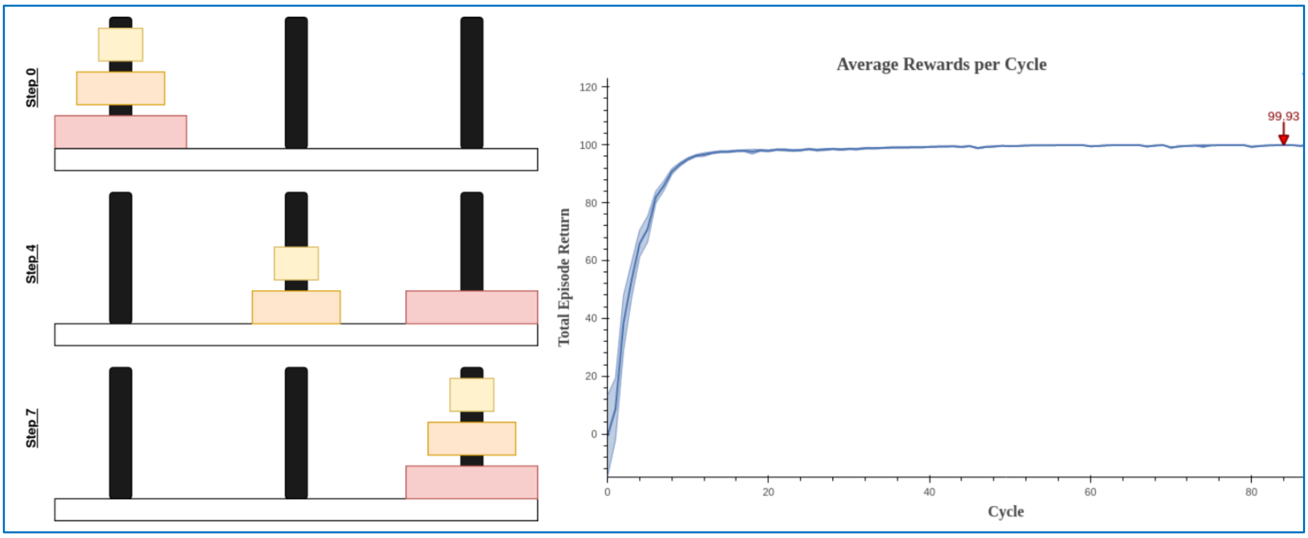

Figure 2: Tower of Hanoi

Left: State of the Tower of Hanoi environment at several steps. Right: Average reward per cycle for an AngoraPy-trained model on the Tower of Hanoi task. The model converges to an optimal solution in about 20 training cycles.

3.3 Publications

Weidler T, Goebel R and Senden M (2023). AngoraPy: A python toolkit for modeling anthropomorphic goal-driven sensorimotor systems. Front. Neuroinform. 17:1223687. doi: 10.3389/fninf.2023.1223687

Weidler, T., & Senden, M. (2023). AngoraPy - Anthropomorphic Goal-Oriented Robotic Control for Neuroscientific Modeling (Version 0.9.0). doi:10.5281/zenodo.6636482

4. Relations to EBRAINS

The work informs the HBP's EBRAINS research infrastructure, in terms of specifications. The Neurorobotics Platform (NRP, SC4) has opted for a modular, containerised design. This was motivated in part by the requirements of the Saccades for Scene Understanding (SSU) architecture. In the course of our work, we actively exchanged ideas on optimal design principles with the NRP developers, driving the evolution of the NRP in a direction that aligns with the modular-integrative approach we developed. Conversely, the work relies on EBRAINS. Specifically, the SSU architecture utilises the NEST simulator for the simulation of a spiking neuron model of saccade generation. Looking Forward

Looking ahead, our aim is to enhance the functionality and efficacy of the modular-integrative modelling framework and AngoraPy. This includes optimising the interface between these tools and users and building a community of contributors. By fostering an inclusive community and creating robust, scalable tools, we aim to drive considerable progress in our understanding of the brain's complex operations.