

Avaliação e Adaptação de um Subsistema de Memória Episódica para uma Arquitetura Cognitiva

K. C. B. Lima E. L. Colombini R. R. Gudwin

Relatório Técnico - IC-PFG-26-36

Projeto Final de Graduação

2026 - Junho

UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE COMPUTAÇÃO

The contents of this report are the sole responsibility of the authors.
O conteúdo deste relatório é de única responsabilidade dos autores.

Avaliação e Adaptação de um Subsistema de Memória Episódica para uma Arquitetura Cognitiva

Karen Cristina Baliero Lima ^{*} Esther Luna Colombini[†]

Ricardo Ribeiro Gudwin[‡]

Resumo

Este trabalho é um relatório final de um projeto realizado como Iniciação Científica, visando a avaliação de um subsistema de memória episódica e sua adaptação para uma arquitetura cognitiva utilizando o CST (Cognitive Systems Toolkit). A realização deste trabalho contou com o suporte financeiro e foi facilitada pela infraestrutura de pesquisa do Hub de Inteligência Artificial e Arquiteturas Cognitivas (H.IAAC) e do Laboratório de Controle e Automação (LCA) instalados no Campus da UNICAMP.

Uma memória episódica é uma funcionalidade essencial na construção de arquiteturas cognitivas, pois permite que agentes aprendam com experiências passadas, melhorando a tomada de decisão e o planejamento futuros. No processo de desenvolvimento de uma memória episódica para o CST, inspirou-se no subsistema de memória episódica desenvolvido pelo CINVESTAV, estudou-se o modelo arquitetural e reproduziram-se os testes no mesmo simulador (Unity) utilizado no projeto original.

Adaptou-se, posteriormente, para a estrutura principal do *CST*: *Memorys* para armazenar informações; *Codelets* intencionais realizando ações paralelas; e *Ideas* como forma de representação do conhecimento. Foi possível reproduzir alguns testes e foram identificadas limitações do modelo original (feito pelo CINVESTAV) que necessitam de aprimoramento antes de disponibilizarmos como uma ferramenta integrada ao CST.

Keywords – Arquiteturas Cognitivas, Memória Episódica, CST

^{*}Instituto de Computação, Universidade Estadual de Campinas, 13081-970 Campinas, SP. Projeto apoiado pelo Ministério da Ciência, Tecnologia e Inovações, com recursos da Lei nº 8.248, de 23 de outubro de 1991

[†]IC - UNICAMP

[‡]FEEC - UNICAMP

Sumário

1	Introdução	3
2	Tipos de memória	3
3	Memória Episódica	5
4	Justificativa	7
5	Objetivos	7
6	Desenvolvimento do Trabalho	8
7	Resultados	16
7.1	Limitações identificadas	16
8	Conclusões	18
8.1	Impactos da Pesquisa	18
9	Produção científica	18
10	Perspectivas de continuidade ou desdobramento do trabalho	19
11	Apoio	19

1 Introdução

Arquiteturas Cognitivas são arquiteturas computacionais inspiradas em modelos cognitivos, que visam reproduzir diferentes capacidades cognitivas humanas em agentes inteligentes [1]. Seu papel principal no contexto da Inteligência Artificial (IA) é permitir a concretização de sistemas artificiais capazes de exibir comportamento inteligente em um ambiente geral, através de uma analogia detalhada com os funcionamentos constitutivos e de desenvolvimento subjacentes à cognição humana [2].

Em [1] são apresentadas informações sobre mais de 900 projetos práticos implementados usando arquiteturas cognitivas nos campos de: modelagem de desempenho humano (*HPM*), jogos e quebra-cabeças, robótica, experimentos psicológicos, processamento de linguagem natural (*NLP*), interação humano-robô e humano-computador (*HRI/HCI*), visão computacional, categorização e agrupamento, agentes virtuais, entre outros. A maioria dessas aplicações, contudo, ainda se concentra no âmbito acadêmico, com menos exemplos na indústria ou em usos sociais.

Historicamente, algumas arquiteturas cognitivas computacionais de propósito geral foram disponibilizadas para uso comum, tais como ACT-R, SOAR [3], CLARION [4], LIDA [?], dentre outras. Apesar de estas arquiteturas estarem disponíveis e, muitas vezes, terem seu código aberto, a grande maioria delas é distribuída na forma de frameworks fechados, que não permitem expansão nem a incorporação de novos métodos e algoritmos em seu funcionamento. Nesse contexto, em 2013, surge o *Cognitive Systems Toolkit (CST)*, um toolkit aberto e expansível para a construção de arquiteturas cognitivas. Ele se baseia em um conjunto de conceitos arquiteturais gerais (chamado de núcleo do *CST*) que permite a composição de novos módulos interligados, com a definição de uma interface comum e genérica. Essa estrutura envolve conceitos familiares a diversas arquiteturas cognitivas, permitindo a integração de partes de outras arquiteturas em seus módulos funcionais [19].

Esse toolkit teve seu desenvolvimento iniciado a partir de um projeto CEPID da FAPESP (CEPID-BRAINN - Proc. 2013/07559-3) e tem sido, desde então, expandido por diversas contribuições de projetos de pesquisa no DCA-FEEC-UNICAMP. Novas ferramentas têm sido desenvolvidas pelo grupo, incorporando diferentes funções cognitivas. Dentre as contribuições, destacam-se o desenvolvimento de funções executivas para aprendizado e tomada de decisão [5], sistemas motivacionais [6], aprendizado instrumental [7], aprendizado sensório-motor [8], ferramentas de visualização e depuração [9], sistemas de memória episódica [11, 12] e aprendizado incremental para robótica [10].

Entretanto, o CST ainda peca na falta de uma memória episódica, assim esse projeto se propõe a estudar e avaliar o modelo de memória episódica desenvolvido em [18], uma arquitetura bio-inspirada, baseada em modelos de diferentes partes do cérebro, desenvolvida pelo CINVESTAV, no México e adapta-la a Arquitetura Cognitiva utilizando o CST.

2 Tipos de memória

Psicologicamente, atribui-se aos humanos a posse de muitos tipos diferentes de memória: memória de curto e longo prazo, memória de trabalho, memória semântica e episódica, entre

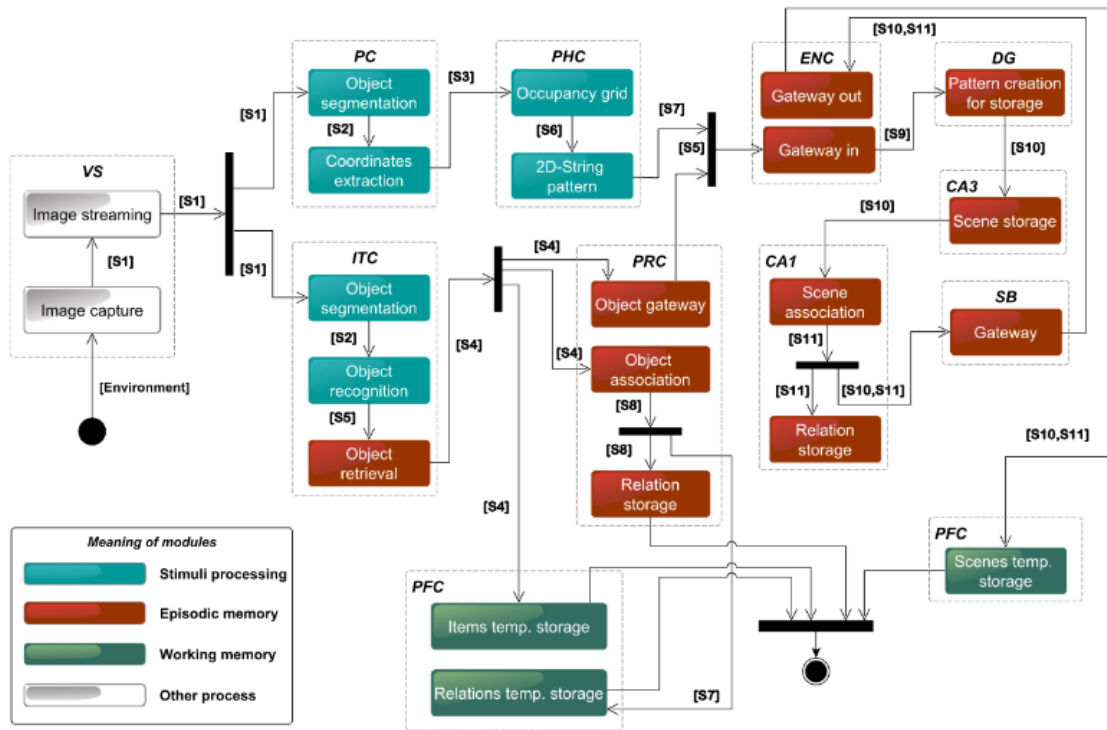


Figura 1: Reprodução do diagrama da arquitetura no texto original de [18]

outras. Embora a pesquisa sobre memória, em diferentes áreas, varie em seus objetivos e perspectivas, elas geralmente consideram o sistema de memória dividido nos seguintes aspectos básicos [13, 14]:

1. Memória Sensorial:
 - a. Memória Icônica;
 - b. Memória Ecoica;
2. Memória de Curto Prazo:
 - a. Esboço Visuoespacial;
 - b. Alça Fonológica;
 - c. Buffer Episódico;
3. Memória de Longo Prazo:
 - a. Memória Não Declarativa:
 - i. Memória Perceptiva;
 - ii. Memória Procedural;
 - b. Memória Declarativa:
 - i. Memória Semântica;
 - ii. Memória Episódica.

3 Memória Episódica

A pesquisa nos tem mostrado que a 'memória' de um sistema cognitivo capacita o sistema a aprender com o ambiente e, mais importante, com a experiência passada [15]. A memória episódica foi proposta por Tulving e descrita por ele como uma verdadeira maravilha da natureza. Em contraste com o aprendizado semântico, o aprendizado episódico lembra eventos e histórias que estão incorporados na experiência. O propósito de um sistema de Memória Episódica é permitir o aprendizado de episódios, ou registros sequenciados temporalmente de eventos específicos que ocorreram com o agente cognitivo [16].

A memória episódica possui algumas propriedades únicas e interessantes. Ela é acionada automaticamente. É autonoética, ou seja, consciente do tempo subjetivo. A arquitetura também possui algumas características únicas. Tem duração variável e é indexada temporalmente. Além disso, a própria estrutura de recuperação possui características interessantes, como a recordação de eventos temporalmente ou conceitualmente próximos, iniciada por uma pista (gatilho). O tempo também é tratado de forma cíclica e isso ajuda na datação. Além disso, a estrutura de armazenamento é tal que muitos eventos são reconstruídos na recuperação e não durante o armazenamento.

Existem 4 fases principais no aprendizado episódico. Elas são:

1. Codificação - como uma memória é capturada e armazenada.

2. Armazenamento - como uma memória é mantida (Dois conceitos importantes são consolidação e esquecimento).
3. Recuperação - como uma memória é recuperada.
4. Uso - como uma memória é usada para melhorar o comportamento (mas isso se baseia nas capacidades gerais da arquitetura, independente da estrutura da memória episódica [16]).

Isso leva a um espaço de designs possíveis, considerando as implicações de cada escolha na funcionalidade da memória episódica. Proposta por [16], a descrição do espaço de design pode ser:

Codificação

- **Iniciação da codificação:** Quando um episódio é codificado? Isso pode ser determinado deliberadamente pelo agente ou pode ser automático em algum momento específico durante o processamento do agente.
- **Determinação do episódio:** Quais são os conteúdos de um episódio?
- **Seleção de características:** Um subconjunto de características nos episódios armazenados é selecionado para correspondência durante a recuperação?

Armazenamento

- **Estrutura do episódio:** Qual é a estrutura do armazenamento da memória episódica? Isso é importante porque a estrutura influencia a eficiência do armazenamento e da recuperação.
- **Dinâmica do episódio:** O conteúdo ou a organização do armazenamento episódico muda ao longo do tempo, como por meio de esquecimento ou generalização?

Recuperação

- **Iniciação da recuperação:** Quando um episódio é recuperado? A recuperação é iniciada deliberadamente, ou a recuperação espontânea é possível usando a situação atual completa como uma pista?
- **Determinação da pista:** Como a pista é especificada?
- **Recuperação:** Dada uma pista, qual episódio é recuperado?
- **Representação do episódio recuperado:** Quando um episódio é recuperado, como ele é representado no agente?
- **Metadados de recuperação:** Existem metadados adicionais sobre o episódio e sua correspondência com a pista?

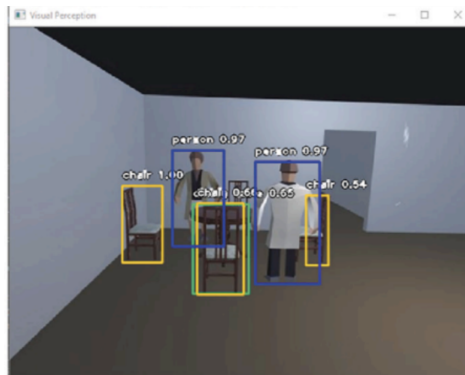


Figura 2: Frame da câmera/perspectiva do agente mostrando o ambiente simulado na Unity em uma das salas com reconhecimento de objetos em *rounding boxes*

Fora da mecânica do sistema de memória episódica, fica a pergunta: "Como a memória recuperada é usada?". Alguns usos possíveis são propostos por [16]:

- **Sensoriamento:** Notar Situações Novas, Detectar Repetição e Sensoriamento Virtual.
- **Raciocínio:** Modelagem de Ações, Modelagem do Ambiente, Previsão de Sucessos/Falhas e Gerenciamento de Objetivos de Longo Prazo.
- **Aprendizado:** Aprendizado Retroativo, Reanálise de Conhecimento, Explicação do Comportamento e "Impulsionar" outros Mecanismos de Aprendizado.

4 Justificativa

A memória episódica tem como importância na inteligência humana o armazenamento de episódios e situações que permitem lembrar acontecimentos do passado, criar relações temporais e associações entre os entes que percebemos. Este trabalho tem como motivação um estudo sobre memória episódica, com a avaliação de um sub-sistema de memória episódica bio-inspirado desenvolvido pelo laboratório CINESTAV. Este projeto realizado no México se propõe a realizar um reconhecimento das experiências de um agente, armazenando esses conhecimentos em cenas, episódios e associações, permitindo os processos de codificação, armazenamento e recuperação que são propostos no design da memória. Na Figura 2 podemos observar uma visão do agente e seu reconhecimento da cena. A partir da adaptação desse projeto para o CST, pretendemos capacitar o toolkit com uma ferramenta de memória episódica a ser utilizada em outros projetos.

5 Objetivos

Nosso objetivo é ter uma ferramenta de memória episódica integrada ao sistema ferramental do CST. Ao realizar a adaptação da memória ao invés de apenas integrá-la, temos como

objetivos secundários:

- Maior facilidade de uso da memória episódica, construída em uma única linguagem (Java)
- Integração com outros projetos que utilizam o CST
- Adaptação do uso da memória episódica para dispositivos móveis

6 Desenvolvimento do Trabalho

O trabalho foi desenvolvido em algumas etapas:

1. Estudo sobre arquiteturas cognitivas e o CST
2. Estudo sobre o sub-sistema de memória episódica do CINVESTAV
3. Execução da arquitetura em ambiente simulado
4. Reprodução dos testes da arquitetura original
5. Adaptação teórica do sub-sistema CINVESTAV para a arquitetura do CST
6. Desenvolvimento do Código no CST
7. Reprodução dos testes na arquitetura CST
8. Análise crítica final acerca da arquitetura CINVESTAV desenvolvida

Etapas 1: Essa pesquisa encontra-se contextualizada na linha de pesquisa "Arquiteturas Cognitivas" ambientada no grupo de pesquisa do H.IAAC (Hub de Inteligência Artificial e Arquiteturas Cognitivas) localizado no IC (Instituto de Computação) e também no DCA (Departamento de Automação Industrial) localizado na FEEC (Faculdade de Engenharia Elétrica e Engenharia da Computação).

O Hub visa o desenvolvimento e disseminação do conhecimento sobre tecnologias capazes de integrar diversos recursos de inteligência em dispositivos móveis, tornando-os hábeis a tomar decisões. Neste contexto, faz-se necessário o estudo sobre o contexto das arquiteturas cognitivas, tendo como base principalmente o artigo de revisão sistemática de [1] além de outros conteúdos necessários na trilha de capacitação para novos pesquisadores no grupo. Foi necessário também um estudo referente ao CST (Cognitive Systems Toolkit) [19], um toolkit destinado à construção de arquiteturas cognitivas, sendo desenvolvido no DCA-FEEC-UNICAMP.

Etapas 2: Na segunda etapa foi feito um estudo sobre memórias episódicas, dentre os diversos modelos de memória episódica, neste projeto iremos investigar o modelo desenvolvido por [18], uma arquitetura bio-inspirada, baseada em modelos de diversas partes do cérebro, desenvolvido pelo CINVESTAV, no México. Esse projeto, que é descrito em detalhes no artigo [18], baseia-se na hipótese de que entidades virtuais, dotadas de uma arquitetura cognitiva dotada de memória episódica, planejarão e tomarão decisões de maneira

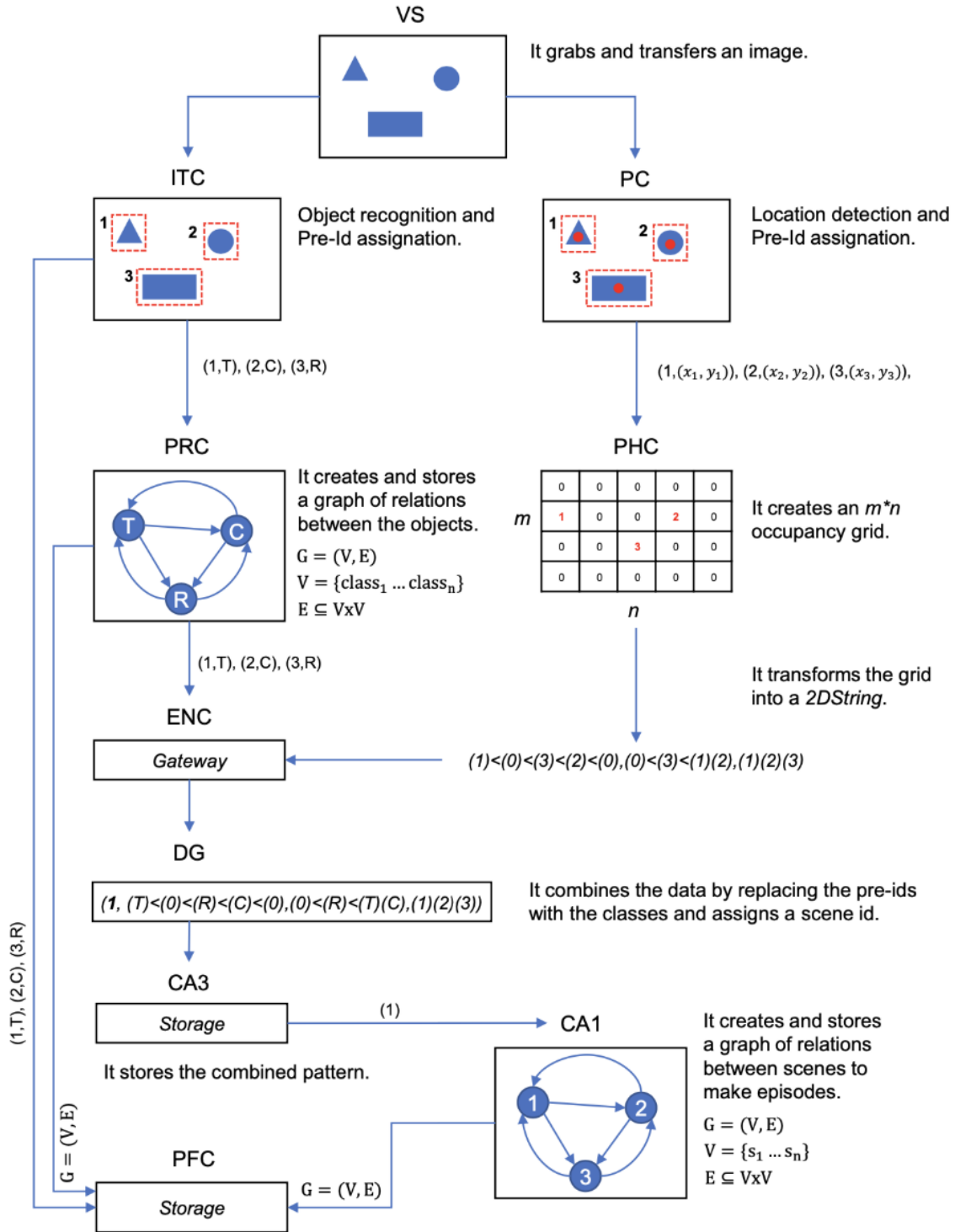


Figura 3: Sumarização do funcionamento da arquitetura

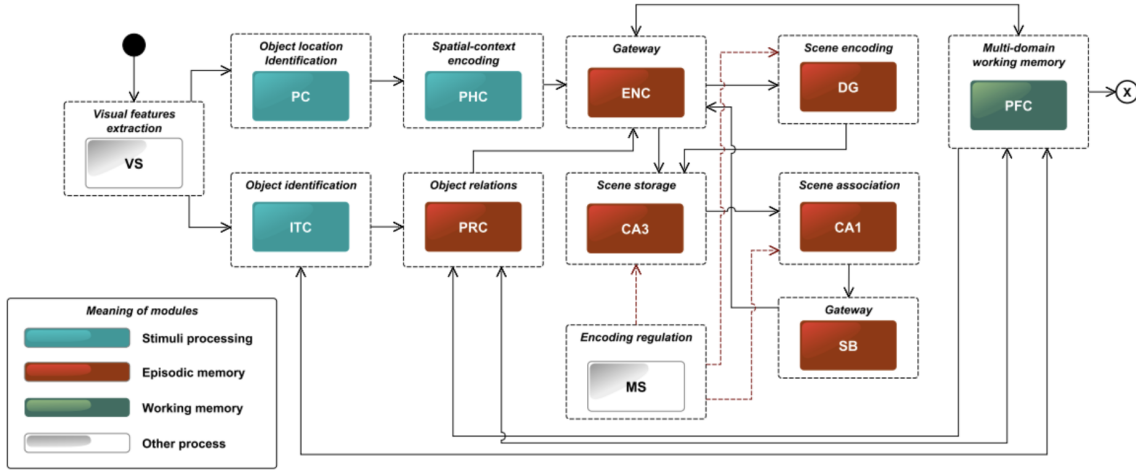


Figura 4: Estruturas cerebrais implicaram em processos de memória episódica [18]

semelhante à humana, utilizando as recordações armazenadas em sua memória episódica. Uma visão sumarizada da arquitetura é mostrada na figura 1.

A descrição a seguir é um pequeno resumo das etapas de processamento sumarizadas na figura 4. Alternativamente, pode-se entender a lógica através da Figura 3.

A primeira parte (VS), que está em cinza na Figura 1, simula o sistema visual do ser humano: as imagens que seriam a visão do agente são capturadas a partir do ambiente simulado e salvas como *byte-arrays* de imagens codificadas e enviadas para o PC e ITC, constituindo a primeira etapa das vias dorsal (para processamento espacial) e ventral (para processamento de características visuais de objetos) do cérebro [18].

No caso da via dorsal a primeira etapa é o PC que representa o Córtex parietal do cérebro. Essa etapa está associada a múltiplas funções relacionados ao processamento espacial (continuando assim a via dorsal) [18]. Nessa parte é realizada a segmentação dos objetos e a extração de coordenadas. Assim é possível saber em quais posições da imagem 2D existem objetos (ainda que não se saiba ainda o que é cada um desses objetos).

Em seguida, é enviado ao PHC que representa o Córtex para-hipocampal. Este realiza o processamento visuoespacial relacionado à percepção da cena, a representação espacial (egocêntrica e aloecêntrica) e a navegação [18]. Nessa parte é criada a *occupancy grid* (para reduzir a complexidade de uma cena visual, o sistema usa uma grade de $M \times N$ linhas e colunas para particioná-la [18]). Na figura 5 é possível observar uma representação da *occupancy grid*.

Posteriormente, a *occupancy-grid* é codificada no formato 2D-String [22], uma versão simbólica dessa matriz. Na Figura 6 é possível observar um exemplo de um *occupancy grid* com algumas posições ocupadas por objetos. Varrendo-se essa matriz, primeiro no sentido vertical e depois horizontal, transforma-se a *grid* especificada em uma 2D-String [18].

O vocabulário de objetos é $\Sigma = \{a, b, c, d\}$, e a matriz da Figura 6 representada por uma

$$G = \begin{bmatrix} z_{11} & z_{12} & z_{13} & \cdots & z_{1n} \\ z_{21} & z_{22} & z_{23} & \cdots & z_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{m1} & z_{m2} & z_{m3} & \cdots & z_{mn} \end{bmatrix}$$

Figura 5: Estruturas cerebrais implicaram em processos de memória episódica [18] pag. 31

d		
	b	c
a	a	

Figura 6: Uma imagem de uma grid/matriz de objetos não reconhecidos

2D-String seria a seguinte:

$$(a = d < a = b < c, a = a < b = c < d)$$

Já pela via Ventral, a primeira etapa é o ITC que representa o Córtex temporal inferior. A área temporal inferior é considerada o estágio final da via de processamento visual ventral e é portanto, bem adequada para processar e representar informações sobre objetos visuais [18].

Esta área é mais orientada ao puro processamento visual, deixando de lado os espaciais [18]. Nessa etapa é realizada a extração de *features*. Em seguida, é realizada a segmentação de objetos e seu reconhecimento, permitindo assim identificar quais são as categorias dos objetos que estão na imagem. O resultado é armazenado e será usado na próxima etapa, o PRC.

Nessa etapa, também é realizada a busca na memória de médio prazo que será usada nas tarefas dos experimentos.

A próxima etapa é o PRC que representa o Córtex perirrinal. Este codifica posições serializadas para cada objeto e gera relações de itens dentro e entre domínios. O córtex perirrinal está envolvido em respostas de familiaridade, que são altamente dependentes da informação do objeto de ligação [18]. Familiaridade representa a consciência de um item ter sido apresentado anteriormente, sem acesso a informações contextuais adicionais sobre o evento [18]. Nesta etapa é realizada a associação entre objetos. A intenção é que a cada dois objetos numa mesma imagem, seu grau de associação aumenta (essas informações são importante para o conceito de familiaridade, assim ao ver um objeto poderia lembrar de outro que possui alto grau de associação com este). Essas informações serão armazenadas a seguir. Nesta etapa é possível realizar busca na memória de médio prazo e recuperar as

associações dos objetos (a serem usadas nos experimentos).

Na próxima etapa temos o ENC que representa o córtex entorrinal. Esta área funciona como um *gateway* de dados. Todos os dados recebidos nesta área são retransmitidos para o destino desejado [18]. Ele armazena os dados para serem usados pelo MPFC e pelo DG.

As informações são enviadas ao DG, este representa o Giro denteado. Ele junta as informações do PHC e PRC criando uma **cena**, em que temos as informações de posição e categoria de cada objeto de uma imagem. As informações recebem um ID exclusivo para que possam ser armazenadas e serão usadas na próxima etapa, o CA3.

O CA3 representa o Corno de Ammon 3, que integra informação espacial e a informação não espacial em representação de objeto-lugar ou evento-contexto. A evidência mostra que o hipocampo desempenha um papel na desambiguação (distinção entre diferentes episódios) de sequências sobrepostas por meio de um processo chamado conclusão de padrão [18].

Este módulo está relacionado aos processos de **armazenamento e recuperação de traços de memória de cenas codificadas** [18]. Nesta etapa é realizada a criação de uma nova *2D-String*, usando os dados vindos da PHC e PRC. Além disso é realizado o armazenamento da cena. Nesta etapa também é realizada a recuperação de cenas similares a uma determinada cena fornecida, processo que será usado nos experimentos. As informações obtidas nessa etapa serão utilizadas na próxima etapa, no CA1.

O CA1 representa o Corno de Ammon 1, que realiza o **armazenamento e recuperação de episódios**. Consideramos um episódio como uma cadeia de múltiplas cenas conectadas [18]. Aqui é realizada a associação entre cenas e seu armazenamento. A associação de cenas estabelece relações binárias entre a cena atual e a anterior. O nível de ativação mostra como as cenas estão relacionadas. Um alto nível de ativação significa que as cenas aparecem em sequência frequentemente [18]. Nesta etapa também é feita a busca de cenas similares na memória de médio e longo prazo para uma dada cena. As informações obtidas nessa etapa serão utilizadas na próxima etapa, no SB.

Esta etapa, SB, representa Subículo que é a principal estrutura de saída do hipocampo. No entanto, não se projeta diretamente para o neocórtex, mas tem conexões de *feedback* com o córtex ENC, que por sua vez se projeta para o neocórtex [23, 25]. Assim como o ENC, esta área funciona como um gateway de dados, ou seja, ela retransmite todos os dados recebidos para o destino desejado [18]. Na prática, ele recebe os dados de CA1 que serão usados pelo ENC.

Existe também a área do PFC (que está em verde na Figura 1) que representa o córtex pré-frontal. Existe uma quantidade considerável de pesquisas bem documentadas sobre o papel do PFC na memória de trabalho (WM). Estudos de danos nesta área sugeriram déficits em pacientes ao executar tarefas na WM. É uma estrutura heterogênea com sub-regiões distintas, como DLPFC, VLPFC e MPFC relacionadas a operações específicas de controle cognitivo [24].

Assim encerramos o estudo sobre o funcionamento do sub-sistema original, e partimos para a próxima etapa de desenvolvimento do código.

Etapa 3: Em parceria do Professor Ricardo Ribeiro Gudwin com o laboratório CIN-VESTAV foi possível obter o código fonte do projeto implementado. Para executá-lo em ambiente simulado foi necessária a configuração do ambiente da máquina local para uso de Unity, Java e Python, além de bibliotecas específicas do YOLO.

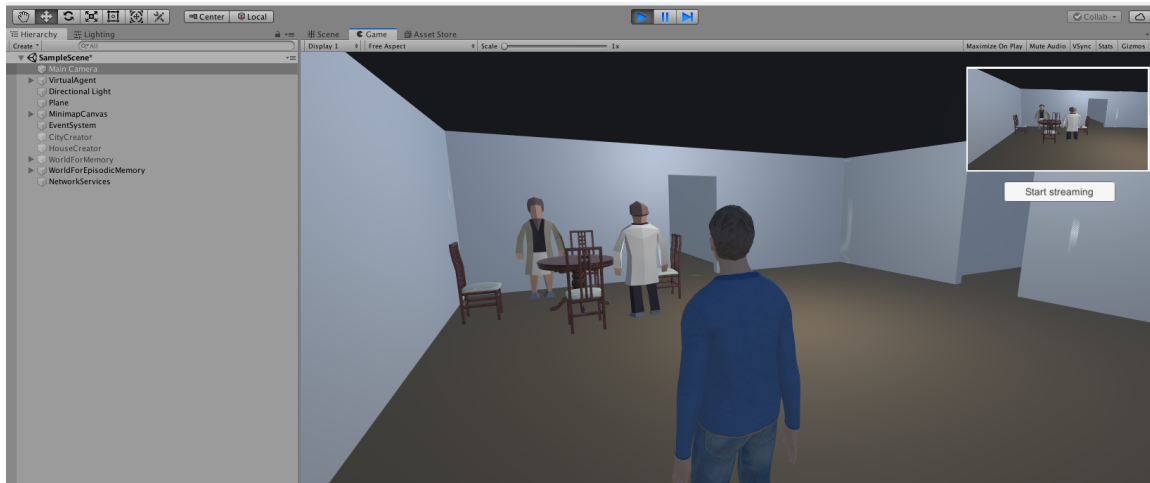


Figura 7: Vista do simulador Unity com o agente/avatar em camiseta comprida a azul e uma sala mostrando duas pessoas e uma mesa marrom que parece ser de madeira

Neste sistema, o processamento é feito através das imagens ou frames recebidas pelo sistema do simulador Unity, e de forma análoga ao que é feito no processamento do cérebro humano, passa a realizar um processamento paralelo das informações, sendo de um lado um processamento voltado a reconhecimento de objetos (identificação de padrões e criação de relações entre objetos) e do outro de localização dos objetivos (espacial).

O ambiente simulado (Unity) contém salas e objetos diferentes, sendo o agente inteligente um avatar controlado pela arquitetura cognitiva, pode-se observar na Figura 7 um exemplo de screenshot da tela da Unity com o avatar em azul.

Etapas 4: Para testar o funcionamento da arquitetura desenvolvida, foram realizadas duas tarefas pelo grupo do CINVESTAV: a primeira tarefa validou as funções da memória de forma independente, e na segunda tarefa, a criatura usou a memória episódica para resolver um problema de planejamento [18]. Essas tarefas foram reproduzidas localmente. Além disso, ambos os experimentos consideram apenas a entrada proveniente do sistema de percepção visual, que é o sentido mais comumente usado em tarefas de memória episódica, e a saída para um sistema de memória de trabalho, que permite a interação com um sistema externo.

- **Experimento 1 (Validação das funções da memória):** Avalia as funções da memória de forma independente: codificação (encoding), armazenamento (storage), recuperação (retrieval) e esquecimento (forgetting) da memória episódica. O agente explora um mundo virtual 3D, adquire conhecimento (aprende o cenário/objetos) e depois sua memória é avaliada através de questões (consultas visuais - pictorial queries) para recuperar associações (como o valor afetivo ou relações espaciais "o que" e "onde"). O esquecimento também é avaliado implicitamente através de como o tempo afeta a recuperação das memórias episódicas.

- **Experimento 2 (Tarefa de Planejamento):** Avalia o uso das capacidades do sistema de memória em uma tarefa específica orientada a um objetivo (goal-directed). O agente é posicionado em um ambiente virtual de vários cômodos (9 salas com objetos diferentes), explora e aprende o mapa/ambiente (estágio de aprendizado). Depois, é instruído a se mover para uma sala/cena específica, exigindo que ele crie um plano de movimento usando o conhecimento adquirido (memórias episódicas).[?]

O trabalho também propõe e realiza experimentos com uso de contexto emocional (reinforcement task), porém não iremos considerar essa parte no presente trabalho.

Em ambos os experimentos, foi possível atingir os objetivos definidos demonstrando o funcionamento da memória episódica.

Etapas 5: Para a adaptação teórica do sub-sistema CINVESTAV deve-se considerar que o modelo arquitetural do CST presume uma separação entre elementos processuais, chamados de *Codelets*, e elementos de armazenagem de dados, chamados de *Memorys*. Essa adaptação foi realizada primeiro pela construção de diagramas, utilizando-se como facilitadora a ferramenta online *draw.io* (um software online gratuito para criar fluxogramas, diagramas de processos, organogramas, diagramas UML, ER e de redes). Além disso, as classes utilizadas no projeto original foram adaptadas para que passassem a ser *Ideas*, forma de representação de conhecimento do CST [20].

Para o desenvolvimento do diagrama utilizou-se o padrão do *CST Core*, observado na Figura 8, que representa os *Memorys* como círculos, e os *Codelets* como retângulos. Um codelet possui duas entradas principais, uma entrada local (In) e uma entrada global (B).

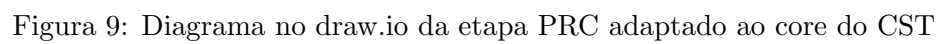
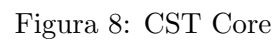
A entrada local é usada para o *codelet* obter informações de objetos de memória, que estão disponíveis na Memória Bruta (*Raw Memory*). A entrada global é usada para o *codelet* obter informações do mecanismo de espaço de trabalho global [21]. As informações provenientes do espaço de trabalho global são variáveis em cada instante de tempo e geralmente estão relacionadas a um resumo, um filtro executivo que seleciona a parte mais relevante das informações disponíveis na memória a cada intervalo de tempo.

Quanto as saídas do *codelet*, temos uma saída padrão (*Out*), que é usada para alterar ou criar novas informações na Raw Memory, e o nível de ativação (*A*), que indica a relevância das informações fornecidas na saída e é utilizado pelo mecanismo de Global Workspace para selecionar informações destinadas ao espaço de trabalho global [19].

Na Figura 9 e na Figura 10 podemos ver como exemplo os diagrama para os casos da etapas do PRC e do DG. As setas representam o fluxo de informações linear entre os codelets e memory objects. Podemos ver dentro de cada *codelet*: o que é executado dentro de seu processamento (representado pelo *proc()*, na Figura 8); além do nome do *codelet*, que na adaptação representará o nome da classe em Java. Abaixo dos *Memorys* temos o seu título, que representa o conteúdo. Na adaptação os *Spikes* do projeto original são colocados como o conteúdo dos *Memorys*. Eles são responsáveis por armazenar o conteúdo produzido pelo processamento de um *Codelet* e podem ser acessados por outros *Codelets*. O diagrama completo pode ser acessado em ¹.

Etapas 6: Na sequência, uma vez que o modelo da arquitetura do CINVESTAV foi compreendido em detalhes, desenvolvemos um código próprio, CST-based, implementando

¹Diagrama completo da arquitetura, disponível no Google Drive



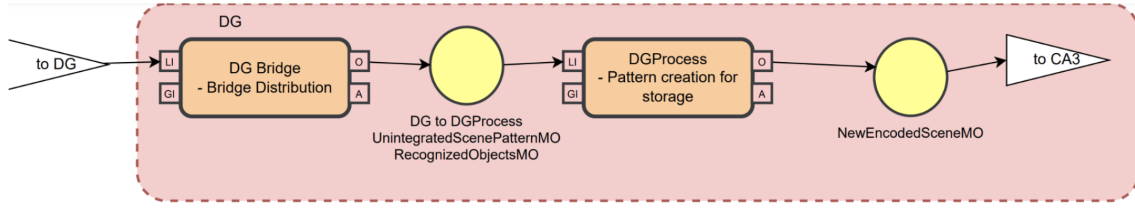


Figura 10: Diagrama no draw.io da etapa DG adaptado ao core do CST

o sub-sistema nas estruturas do CST. O código foi disponibilizado em repositório público no github do H.IAAC². Dentre as vantagens do uso do CST está a possibilidade de visualizar, através do MindViewer, interface de inspeção e visualização das estruturas internas de uma arquitetura cognitiva, o estado atual dos codelets e memory objects, conforme exemplo na Figura 11. Foi necessária também a adaptação do reconhecimento de objetos que era feita com bibliotecas nativas do Python; além disso o projeto original se utiliza da ideia de Spikes, que foram transformados em *Memorys* e *Ideas*.

Etapa 7: Buscou-se reproduzir os mesmos testes feitos originalmente pelo grupo CIN-VESTAV, porém agora com a arquitetura utilizando o *CST*. Conseguimos adaptar o *encoding*, *storage* e *retrieval* das informações, além de acessar as informações acerca da associação de objetos, buscar relações e cenas. Porém não conseguimos acessar os testes feitos para buscar caminhos.

Etapa 8: Foi realizada como última etapa uma análise criteriosa quanto a ferramenta, suas facilidades e dificuldades, para formalizar opinião crítica sobre a incorporação dessa ferramenta no CST.

7 Resultados

Como resultado do projeto, foi feita a adaptação do sub-sistema de memória episódica para o CST, incluindo a adaptação da conexão da arquitetura com o simulador e o reconhecimento de objetos em imagens, agora disponíveis em Java. Também foi feita a adaptação das etapas para *Codelets*, *Memorys* e *Ideas* e uso do *MindViewer* para visualização da execução. Conseguimos visualizar as interfaces dos testes. O projeto está disponibilizado no GitHub do grupo de pesquisa H.IAAC³ e foi criado um script para instalação no ambiente linux.

7.1 Limitações identificadas

- Esse projeto não tem a capacidade de identificar perspectiva da imagem processada, ou seja, objetos que estão sobrepostos ou em distâncias diferentes são identificados da mesma forma uma vez que as imagens processadas são em 2D.

²Repositório do projeto desenvolvido no Github

³<https://github.com/H-IAAC/episodic-memory/tree/develop>

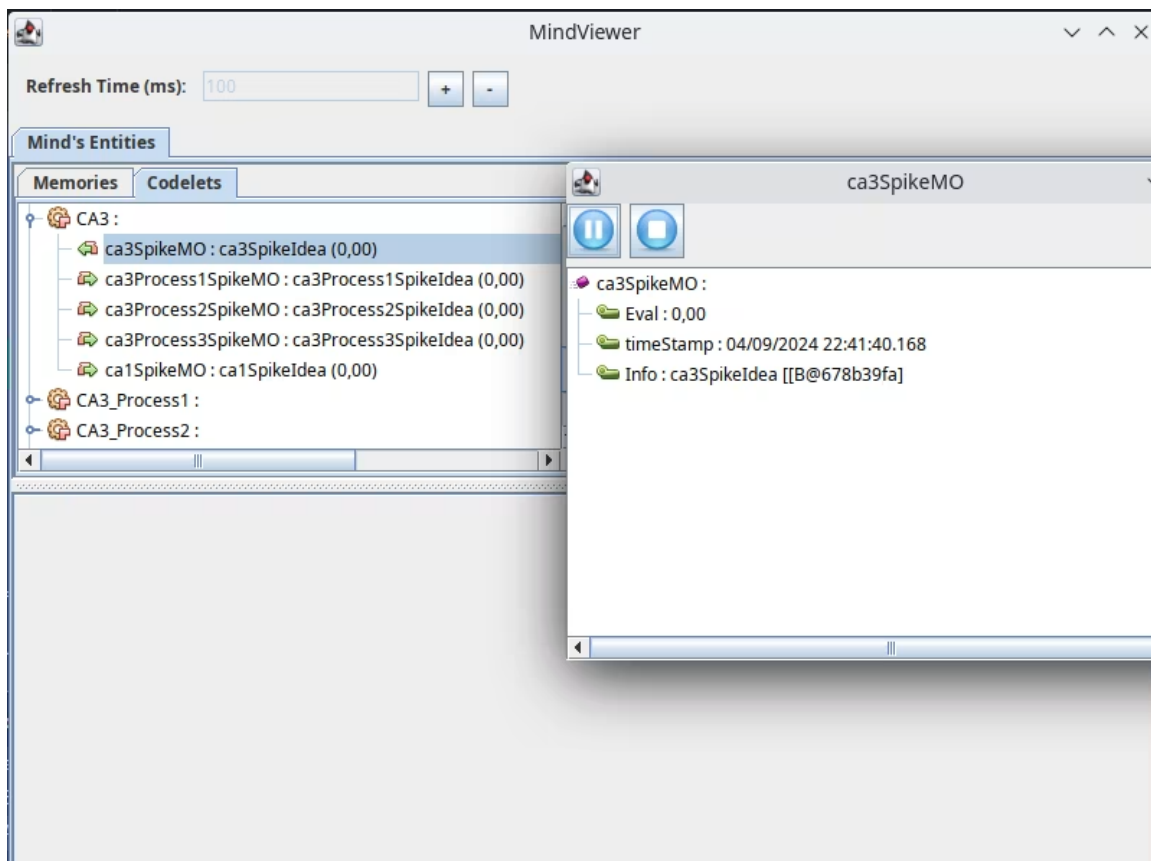


Figura 11: Interface do MindViewer com foco no Memory Object CA3Spike MO do Codelet CA3

- O reconhecimento de objetos é limitado ao número de classes de objetos da biblioteca escolhida. Os objetos reconhecidos pelo agente são limitados pela biblioteca já pré-treinada escolhida pelo programador do sistema, assim haverá falhas em caso de objetos na cena que não estejam dentro do conhecimento do modelo escolhido.
- Falta de identificação de características que diferenciem objetos de uma mesma categoria entre si. Como por exemplo: cores, formas, tamanhos, etc, assim todos os objetos, mesmo que diferentes na prática, serão generalizados em categorias que os agrupam.
- Não há muita margem para aprendizado, uma vez que utiliza-se de modelos pré-treinados para classificação de objetos.
- Falta uma generalização uma vez que o projeto requer o uso de imagens, portanto esse subsistema de memória episódica irá se aplicar a processamento visual apenas.

8 Conclusões

Esperava-se obter resultados similares aos obtidos em [18], porém agora utilizando-se a tecnologia do CST. Isso foi parcialmente concluído uma vez que conseguimos adaptar o *encoding*, *storage* e *retrieval* das informações, além de acessar informações acerca da associação de objetos, buscar relações e cenas. Entretanto, não conseguimos acessar os testes feitos para buscar caminhos. Com a execução deste projeto foi possível concluir parcialmente os objetivos da adaptação da arquitetura com o uso do CST. Como identificado na seção de resultados existem muitos desafios e limitações arquiteturais sobre o projeto desse sub-sistema que devem ser atacados antes de fornecer como ferramenta do CST.

8.1 Impactos da Pesquisa

O principal impacto dessa pesquisa foi o ganho de *expertise* por parte de nosso grupo, de uma metodologia concreta para o desenvolvimento de uma memória episódica e a observação de seu funcionamento. Uma vez que a memória não será incorporada ao CST, seu uso se limitará ao estudo teórico e acadêmico de uma tentativa de reprodução da memória episódica bio-inspirada em máquinas artificiais.

9 Produção científica

As produções científicas derivadas do projeto de pesquisa foram:

- Artigo de divulgação científica realizado na disciplina de pós graduação da FEEC-UNICAMP: IA718G - Tópicos em Sistemas Inteligentes. O artigo "A Literature Review on Episodic Memory for Cognitive Architectures" não foi publicado mas encontra-se neste link⁴.

⁴Artigo "A Literature Review on Episodic Memory for Cognitive Architecture"

- Participação no XVIII Workshop de Teses, Dissertações e Trabalhos de Iniciação Científica (WTD) realizado no Instituto de Computação da UNICAMP. Foram realizadas a produção e apresentação de pôster⁵, apresentação de 10 min Lightning Talk⁶ e criação de um vídeo de 1 minuto disponibilizado no instagram do evento⁷.
- Criação e apresentação de pôster no II Workshop do H.IAAC⁸

10 Perspectivas de continuidade ou desdobramento do trabalho

Modularização: Algumas partes do projeto (desenvolvidas em Codelets) podem ser adaptadas e integradas ao CST, tornando-as reutilizáveis em outras arquiteturas, permitindo o processamento em diferentes contextos.

Reprodutibilidade: Visando a reprodutibilidade, o projeto deverá ser configurado para usar Docker, possibilitando a execução em outros sistemas operacionais sem grandes dificuldades na instalação de requisitos.

Capacidade: Necessidade de criação de um buffer de memória para lidar com processamento em alto volume.

Transferência de conhecimento para outros contextos: Possibilidades de desdobramentos do trabalho incluem a conexão com outros simuladores além da Unity, e até mesmo testes com câmeras capturando o mundo real.

11 Apoio

Este projeto foi apoiado pelo Ministério da Ciência, Tecnologia e Inovações, com recursos da Lei nº 8.248, de 23 de outubro de 1991, no âmbito do PPI-SOFTEX, coordenado pela Softex e publicado Arquitetura Cognitiva (Fase 3), DOU 01245.003479/2024-10

Agradecimentos

Agradeço a todos que me apoiaram durante o projeto e que tem sido compreensivos com minha dedicação a pesquisa. À minha família e amigos por sempre estarem ao meu lado e darem todo o apoio que preciso, permitindo que eu me dedique aos estudos à pesquisa científica. Agradeço ao orientador Ricardo Gudwin, aos demais coordenadores e pesquisadores do H.IAAC pelo apoio, orientação e confiança.

⁵Vídeo no Instagram do WTD

⁶Slides apresentação Lightning Talk WTD

⁷Vídeo no Instagram do WTD

⁸Pôster do workshop H.IAAC

Referências

- [1] Kotseruba, Iuliia and Tsotsos, John K., *A Review of 40 Years of Cognitive Architecture Research: Core Cognitive Abilities and Practical Applications*, arXiv:1610.08602 (2018).
- [2] Lieto, Antonio and Bhatt, Mehul and Oltramari, Alessandro and Vernon, David, *The role of cognitive architectures in general artificial intelligence*, Cognitive Systems Research 48, 1–3 (2018).
- [3] Laird, John E., *Extending the soar cognitive architecture*, Frontiers in Artificial Intelligence and Applications 171, 224–235 (2008).
- [4] Sun, Ron, *The CLARION cognitive architecture: Extending cognitive modeling to social simulation*, Cognition and multi-agent interaction, 79–99 (2006).
- [5] Raizer, Klaus, *Executive functions for Learning and decision-making in a bio-inspired cognitive architecture*, Universidade Estadual de Campinas (2015).
- [6] Froes, Eduardo de Moraes, *Construindo um subsistema motivacional para o cognitive system toolkit*, Universidade Estadual de Campinas (2017).
- [7] Gibaut, Wandemberg Santana Pharaoh, *Uma arquitetura cognitiva para o aprendizado instrumental em agentes inteligentes*, Universidade Estadual de Campinas (2018).
- [8] Rossi, Leonardo de Lellis, *Aprendizado sensório-motor em robôs cognitivos utilizando modelo CST-CONAIM*, Universidade Estadual Paulista (2021).
- [9] Marques, Áureo Henrique e Silva and Raizer, Klaus and Martins, Paulo and Gudwin, Ricardo, *Visualization tools for monitoring and debugging a cognitive architecture using CST*, Procedia Computer Science 213, 528–535 (2022).
- [10] Berto, Letícia Mara, *A motivation-driven incremental learning framework for robotics*, Universidade Estadual de Campinas (2025).
- [11] Sakabe, Eduardo Yuji, *Building an episode tracker module for the cognitive systems toolkit*, Universidade Estadual de Campinas (2024).
- [12] Silva, Bruno Guedes da, *Sistema de memória episódica para agentes inteligentes utilizando CST e representação de conhecimento Idea*, Universidade Estadual de Campinas (2024).
- [13] Tulving, E., Hayman, C. & Macdonald, C. Long-lasting perceptual priming and semantic learning in amnesia: A case experiment.. *Journal Of Experimental Psychology: Learning, Memory, And Cognition*. **17**, 595-617 (1991), <https://doi.org/10.1037/0278-7393.17.4.595>

- [14] Baddeley, A., Della Sala, S., Papagno, C. & Spinnler, H. Dual-task performance in dysexecutive and nondysexecutive patients with a frontal lesion.. *Neuropsychology*. **11**, 187-194 (1997), <https://doi.org/10.1037/0894-4105.11.2.187>
- [15] Kuppuswamy, N., Cho, S. & Kim, J. A Cognitive Control Architecture for an Artificial Creature using Episodic Memory. *2006 SICE-ICASE International Joint Conference*. pp. 3104-3110 (2006)
- [16] Nuxoll, A. & Laird, J. Extending Cognitive Architecture with Episodic Memory.. (2007,1)
- [17] Dodd, W. & Gutierrez, R. The role of episodic memory and emotion in a cognitive robot. *ROMAN 2005. IEEE International Workshop On Robot And Human Interactive Communication, 2005..* pp. 692-697 (2005)
- [18] Martin, Luis et al., *Declarative Working Memory: A Bio-Inspired Cognitive Architecture Proposal*, Cognitive Systems Research, 66 (2020).
- [19] Paraense, Andre LO et al., *The cognitive systems toolkit and the CST reference cognitive architecture*, Biologically Inspired Cognitive Architectures, 17, 32–48 (2016).
- [20] Camargo, Eduardo et al., *Existence, Hypotheses and Categories in Knowledge Representation*, Procedia Computer Science, 213, 496–503 (2022).
- [21] Baars, Bernard J., *Global workspace theory of consciousness: toward a cognitive neuroscience of human experience*, Progress in Brain Research, 150, 45–53 (2005).
- [22] Chang, Shi Kuo and Jungert, Erland, *Symbolic projection for image information retrieval and spatial reasoning*, Academic Press London (1996).
- [23] Witter, Menno P. et al., *Architecture of the entorhinal cortex a review of entorhinal anatomy in rodents with some comparative notes*, Frontiers in Systems Neuroscience, 11 (2017).
- [24] Wolf, Robert Christian et al., *Differential activation of ventrolateral prefrontal cortex during working memory retrieval*, Neuropsychologia, 44(12), 2558–2563 (2006).
- [25] O'Mara, Shane, *Controlling hippocampal output: The central role of subiculum in hippocampal information processing*, Behavioural Brain Research, 174(2), 304–312 (2006).