1 2 9 0

UNIVERSIDADE Ð COIMBRA

Gustavo Miguel Santos Assunção

Emotional Competence and Adaptability in Artificial Intelligence

Tese no âmbito do Doutoramento em Engenharia Electrotécnica e de Computadores, ramo de especialização em Computadores e Electrónica, orientada pelo Professor Doutor Paulo Jorge Carvalho Menezes e pelo Professor Doutor Miguel de Sá e Sousa de Castelo-Branco, apresentada ao Departamento de Engenharia Electrotécnica e de Computadores da Faculdade de Ciências e Tecnologia da Universidade de Coimbra.

Novembro de 2023

1 2 9 0

UNIVERSIDADE D COIMBRA

Faculty of Science and Technology
Department of Electrical and Computer Engineering

Emotional Competence and Adaptability in Artificial Intelligence

Gustavo Miguel Santos Assunção

Thesis submitted to the Department of Electrical and Computer Engineering of the Faculty of Science and Technology of the University of Coimbra in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Supervised by:
Prof. Dr. Paulo Jorge Carvalho Menezes
Prof. Dr. Miguel de Sá e Sousa de Castelo-Branco

Coimbra, 2024

This work was developped in collaboration with:

University of Coimbra

Department of Electrical and Computer Engineering

Institute of Systems and Robotics

Pós: *I am the king of my own land*

AG

Pré: *Eu queria ser astronauta, o meu país não deixou...*

Tim

Agradecimentos

Sem dúvida existe uma dualidade de impacto que pessoas têm nas nossas vidas, tipicamente ignorada ao abrigo de uma ignorância consciente e permissível de agradecimento e apreciação do bom. Nesse aspecto agradeço aos meus pais Dina e Jorge a quem eu sempre quis impressionar e pressentir orgulho em mim, que na sua imcompreensão sempre me incitaram a lutar por isto. Agradeço ao Paulo, o meu orientador que em momentos de desespero se revelou um mentor e amigo. Agradeço ao Bruno, irmão que lutou contra isto lado a lado comigo e me segurou quando eu quase caí. À Joana agradeço por anos de suporte e apoio incondicional, os quais eu valorizo mais do que lhe demonstrei. Ao Samuel que, na voz da razão, me dizia para largar o que me fazia mal. Ao Júnior por compreender o meu pensamento e ao máximo tentar ajudar a estruturá-lo em algo saudável. A estas pessoas agradeço por terem estado lá para mim quando eu precisei e lamento se não consegui ser para eles e elas o que foram para mim.

Finalizando, devo também agradecer ao meu irmão André e cunhada Elif, tal como aos meus tios Teresa e Amílcar, por sempre acreditarem nas minhas capacidades. Também o meu co-orientador Miguel Castelo-Branco, e toda a equipa atual e anterior do IS3L merecem a minha gratidão, bem como o geral do ISR e quem de lá me auxiliou ao longo destes anos. Agradeço também ao Prof. Humberto Jorge por me emprestar o seu hábito talar para a minha defesa. Tenho também de agradecer à equipa do Overleaf e ao criador do LaTeX, sem os quais naturalmente este documento não existiria. Compactuando também com o Bruno, acabo com um agradecimento aos iogurtes Skyr (e proteicos em geral) por terem sido um combustível indispensável.

De resto, como autor agradeço à Universidade de Coimbra, especificamente ao Departamento de Engenharia Electrotécnica e de Computadores e ao Instituto de Sistemas e Robótica por me providenciarem todas as condições necessárias e recursos para poder atingir as metas estabelecidas. Este trabalho foi também financiado pela Fundação para a Ciência e a Tecnologia (FCT), sob tutela do Ministéria da Ciência, Tecnologia e Ensino Superior (MCTES) do Governo de Portugal, com a bolsa de referência 2020.05620.BD. Os projectos de referência UIDB/00048/2020 e UIDP/EEA/00048/2020 também suportaram este trabalho ocasionalmente.

Abstract

The advent of deep learning and Artificial Intelligence (AI), combined with the recent hardware advancements that made it possible, has brought on a wave of new and successful approaches to many problems previously unsolvable. However, critical challenges remain linked with algorithm dependency on big data, user input, and related issues, hindering autonomy. Simultaneously, increasingly independent systems must account for user admissibility as well as develop rapport with the human counterparts they accompany. These issues require attention if the current progress rate in AI is to continue. Thus, the main focus of this research was the study and development of deep learning solutions pertaining to emotional competence and adaptability, emulating biological factors that mitigate the specified issues. Consequently, this document encompasses two major parts where the developed solutions are appropriately presented within each of their corresponding areas of research.

While formulation and development of bio-inspired solutions encompassed most of this thesis' duration, such an endeavor was premised by familiarization with and computational adaptation of how biological phenomena occur and affect the brain concerning learning, affection, and information processing. This constitutes the first part of the manuscript and was done to ensure the validity of the solutions with respect to Neurophysiology. Respectively, neural circuitry involved in emotional processing and demeanor adaptation to stimuli was overviewed in depth. This was studied in tandem with the interpretative take of Psychology on the effect emotion has on decision-making and behavior. Principal factors to consider for both emotion recognition and expression by artificial agents were also surveyed. Moreover, adaptive characteristics, particularly relating to neuromodulation and plastic re-organization, were assessed in terms of their suitability for emulation within the context of artificial intelligence.

Solutions within the context of emotion recognition/expression were addressed in the second part of this thesis, mitigating issues on the loss of valuable information and the development of empathic behavior by agents. Both resulted in successful improvements within their respective fields, with applications envisioning the service and assistive systems as well as companionship robotics largely for health and eldercare. Regarding the third part of this thesis, two different methods were proposed following dopamine emulation in artificial neural networks. While effects over learning efficiency were negligible, hindering the initial goal, an interesting parallelism was observed between biological neuromodulation and this work. This is particularly important as it still demonstrates the validity of the work developed. Secondly, given interdisciplinary knowledge's importance for the advancement of research, an experiment was designed based on psychology guidelines for human participants and a neurological background implemented for correlating emotional stimuli with action selection. With this foundation, artificial agents successfully demonstrated surprise-exploration and pride-exploration correlations similar to those learned by living beings for adequate survival and goal achievement. Naturally learning useful correlations between intrinsic qualities and decision-making is tremendously impactful for AI, with the outcome of this research having spun off into works on more autonomous social robotics. The final part of this thesis presented open-ended questions regarding the previous chapters as well as proposed a bio-inspired strategy for dealing with overfitting issues prevalent in neural network training, with promising early-stage results.

Keywords

Adaptability; Artificial Intelligence; Deep Learning; Emotional Competence; Reinforcement Learning

Resumo

O crescimento de *deep learning* e AI, combinado com os recentes avanços em hardware que o permitiram, despoletaram uma onda de novas e bem-sucedidas abordagens a vários problemas, antes irresolúveis. No entanto, diversos desafios existem ainda relacionados com dependência de algoritmos em *big data*, *input* do utilizador e outros problemas impeditivos de maior autonomia. Simultaneamente, sistemas mais independentes devem ter em consideração a sua aceitação por parte do utilizador, bem como o desenvolvimento de conexão com os humanos que acompanham. Estes problemas requerem especial atenção, para que esta taxa de progresso em AI se mantenha. Como tal, o cerne desta tese foi o estudo e desenvolvimento de soluções em *deep learning* ligadas à competência emocional e adaptabilidade, através da emulação de fatores biológicos mitigantes dos problemas mencionados. Consequentemente, este documento inclui duas partes principais e distintas, onde as soluções desenvolvidas são apresentadas dentro das suas áreas de pesquisa correspondentes.

A formulação e desenvolvimento de soluções bio-inspiradas ocupou grande parte da duração desta tese. Contudo, tal foi precedido por familiarização e adaptação para computação dos fenómenos biológicos que ocorrem e afetam o cérebro no que toca a aprendizagem, afetividade e processamento de informação. Este segmento de trabalho constitui a primeira parte do manuscrito e foi feito de modo a assegurar a validade das soluções relativamente a Neurofisiologia. Respetivamente, circuitos neuronais envolvidos em processamento emocional e adaptação comportamental a estímulos foram estudados compreensivamente. Conjuntamente foi explorada a perspetiva interpretativa da Psicologia sobre os efeitos que emoção acarreta sobre comportamento e tomada de decisões. Deste modo, foram averiguados quais os principais fatores a considerar tanto para reconhecimento como para expressão de emoção por parte de agentes artificiais. Adicionalmente, caraterísticas adaptativas, maioritariamente relacionadas com neuromodulação e re-organização plástica, foram avaliadas em termos da sua adequação à emulação no contexto de inteligência artificial.

Trabalhos enquadrados na área de reconhecimento/expressão de emoção foram expostos na segunda parte desta tese, com um intuito de mitigar problemas relacionados com perda de informação valiosa e desenvolvimento de empatia por agentes. Ambos os trabalhos foram bem-sucedidos e resultaram em melhorias dentro das suas respetivas áreas, com aplicações ponderadas para sistemas de assistência e serviço, bem como robótica de companhia no campo da saúde e tratamento de idosos. Em relação à terceira parte desta tese, dois métodos diferentes foram propostos em concordância com emulação de dopamina em redes neuronais artificiais. Enquanto que efeitos sobre eficiência na aprendizagem foram desprezíveis, travando o cumprimento do objectivo inicial, um paralelismo interessante foi observado en-

tre este trabalho e neuromodulação biológica. Isto é particularmente importante uma vez que contribui para a validação do trabalho desenvolvido. Além deste trabalho, visando a importância do conhecimento interdisciplinar para o avanço de pesquisa, uma experiência foi desenhada tendo por base orientações da psicologia explicitadas para participantes humanos e também numa base neurológica responsável por correlacionar estímulo emocional com seleção de ações. Estabelecidos estes alicerces, agentes artificiais demonstraram com sucesso correlações surpresa-exploração e orgulho-exploração semelhantes às adquiridas por seres vivos, para sobrevivência e cumprimento de objetivos. Naturalmente, aprendizagem de correlações úteis entre qualidades intrínsecas e tomada de decisões é extremamente impactante para AI, com o resultado desta pesquisa tendo levado a outros trabalhos de aplicação em robótica social mais autónoma. A parte final desta tese apresentou questões abertas com respeito aos capítulos anteriores, assim como propôs uma estratégia bio-inspirada para lidar com questões de *overfitting* prevalentes no treino de redes neuronais. Resultados iniciais desta abordagem demonstraram ser promissores.

Palavras-Chave

Adaptabilidade; Aprendizagem Profunda; Aprendizagem por Reforço; Competência Emocional; Inteligência Artificial

Contents

List of Figures

List of Tables

List of Algorithms

Acronyms

ACC anterior cingulate cortex

AGI general AI

AI Artificial Intelligence

ANN Artificial Neural Network

ASR automatic speech recognition

BG basal ganglia

CNN Convolutional Neural Network

DDPG Deep Deterministic Policy Gradient

DL deep learning

DPG Deterministic Policy Gradient

DQN Deep Q-Learning Network

DS Dorsal Striatum

FER facial emotion recognition

HRI human-robot interaction

MDP Markov Decision Process

ML machine learning

mPFC medial prefrontal cortex

NAc nucleus accumbens

OFC orbitofrontal cortex

PE Prediction Error

PFC prefrontal cortex

RL reinforcement learning

RNN Recurrent Neural Network

SER speech emotion recognition

SL supervised learning

SN substantia nigra

STDP Spike Timing Dependent Plasticity

TD temporal difference

VS Ventral Striatum

VTA ventral tegmental area

Chapter 1

Introduction

Contents

This thesis addresses AI and deep learning (DL) specifically as areas which, despite fomented by an intent to emulate attributes of a brain to solve complex computational tasks, have deviated from a biological basis. Given its potential, a bio-inspired methodology, largely but not solely related to emotion, is proposed and its performance is reported to identify viable paths for future DL research.

1.1 Context

AI is an umbrella term for decision-making and task-oriented methodology, seemingly resembling human cognition/behavior, which is employed in computers and other machinery. Specifically, typical machine learning (ML) procedures involve training a model to perform some task the user intends to automate, which is employed once performance metrics reach acceptable values. Inputs may be sourced from virtually any modality, through corresponding peripherals (e.g. microphones, RGB-D cameras, various sensors). As the integration of AI into industry and society becomes more prominent, new challenges emerge and prompt an increased focus on innovation. If solutions are to be widely applicable and not single-use, standard goals such as adaptability, admissibility by users, and to some degree autonomy, become pivotal in research. Naturally, innovations must be developed under simulated conditions from which knowledge can be obtained and transferred to real-world deployment. This is necessary to account for the lack of data or its difficult collection in several areas, end-user safety, and also the strain it would put on real-world platforms to serve as test benches of novel techniques.

Biological neural functioning provides a multitude of unmatched examples of how cognition can be developed and improved. For instance, the neurotransmission processes linking the cerebral cortex with limbic structures such as the amygdala demonstrate how emotion can mediate learning by instigating internal drives [1, 2]. Relatedly, emotional competence in humans has been flagged as paramount for adequate development as well as social success [3]. Other examples include neuromodulatory processes and plasticity, which confer to the brain an ability to optimize paths for specific tasks, by potentiating frequently activated neuron connections and weakening underused ones [4]. Also, neural pattern replay during dreamless sleep has been shown critical for abstracting core knowledge and consolidating memory [5]. Despite being dismissed by most AI research, emulation of these and other useful characteristics of brain operation could prove highly beneficial for the field.

DL is a particular branch of ML and AI whose performances have progressively surpassed those of conventional methods and which can largely benefit from the emulation of biological processes. Artificial Neural Networks (ANNs), whose units model real neurons, could very well be improved in terms of adaptability were internal drives and plasticity integrated into their functioning. AI agents would benefit from emotional competence in terms of user acceptance. Considering these and other cases, it seems worthwhile to explore further emulation of biological neural processes in DL.

1.2 Motivation

Longevity has been demonstrated to be a preponderant issue with implications towards sustainable development in this day and age [6]. A decay in health and overall life quality matched by failure to accommodate caregiving demands are natural consequences of this development path. Complications encompass both physical aspects and psychological factors, as people experience a growing body of disorders plus reduced independence, along with loneliness or lack of companionship. This scenario unfortunately worsens daily and requires innovations in healthcare technology so as not to spiral down further. In line with aging, statistical data points to the shrinkage of the active population as another obstacle to sustainability. This reduction has a socioeconomic impact with a wider reach than mere assistive services. As several activities in industry and infrastructure depend on manual intervention to function properly, the growing lack of humanpower also becomes a time-sensitive issue.

Some recent technological advances in the fields of robotics and AI have provided leeway in terms of meeting the demands that aging and shrinkage of active population have created. Nursing and service robotics are demonstrative of this [7] but still show limitations concerning user acceptance, adaptability, navigation, and other factors that prevent wider adoption. Likewise, identical problems are prevalent in industry-adopted AI and ML solutions [8], particularly pertaining to autonomy and requirements related to putting or maintaining the *human in the loop*. These societal challenges could be mitigated to some degree by biologically emulating DL, as it can provide a range of solutions that are not currently available. To exemplify, emotionally competent agents can integrate socially and offer care better suited to users. Additionally, internally driven models capable of restructuring are better equipped to deal with unexpected scenarios or new tasks.

Besides assistive or industrial machinery, it would be senseless to not consider other consumer-level robotics as tremendous beneficiaries of bio-emulating DL. Particularly emotional competence can spur a new generation of socially interactive robots, requiring engagement with users to be perceptively affective [9] for easy acceptance. Thus, the implementation of procedures and/or limbic structures associated with the perception, expression, or learning-related aspects of emotion most likely will become a standard for human-robot interaction (HRI) shortly, adding to the importance of their study now.

On a side note from robotics, neuroscientific and psychological findings tend to be analyzed separately [10], despite often describing the same processes. Moreover, there is a lack of frameworks on which theories of neural functioning can be corroborated. This issue ultimately stunts research, as it prevents the correlation between interdisciplinary findings and the consequential advancement of the fields. Nevertheless, DL emulating neural processes could be used to resolve simulations for both Neuroscience and Psychology. By recreating systems as DL models operating according to a proposed theory, experimental results can serve as empirical proof of that theory's veracity or lack thereof. As a concrete example, one could perform modifications to an ANN in a manner thought to be related to developmental disorders in a real brain, and observe its later performance. This is still limited to basic dynamics, such as [11] observing habituation patterns in emotional reinforcement learning (RL) agents.

Given these reasons, DL research should certainly prioritize analyzing the benefits of biological neural processes and consider integration when adequate. Hence this thesis focuses on this very topic, presenting a detailed evaluation on DL implementations regarding emotional competence and adaptability in terms of interaction, learning, and other scenarios. Validation is performed on standard datasets or real-world scenarios when possible. Contributions developed during the duration of this doctoral program are addressed in the following sections.

1.3 Primary Contributions

The main objectives of this thesis encompassed presenting techniques useful for DL adaptability and learning autonomy, as well as user acceptance and engagement, largely focusing on applications within assistive robotics and general AI research. Here are listed the main themes approached during its duration.

- **Emotional competence for improved assistive services and HRI** - A set of techniques for emotion recognition, assessment, and expression focusing on better understanding user affective needs and evolving robotic agents, from tools to companions. The recognition and expression of emotion were explored in the auditory and visual modalities, interfacing via standard cameras or microphones and computer screens or basic LED arrays, as these are commonly available and can be easily deployed in most cases. The evaluation involved elements of supervised learning (SL), fuzzy logic, and RL, either developed solo or combined for application in two devised experimental scenarios. Specifically, one of these performed fuzzification of emotional features extracted from audio to benefit from the correlation of absolute states and improve recognition rates. The other explored continuous optimization of facial emotion expression using deep Q-networks and user feedback for later empathic interactions post-user emotion recognition.

- **Dopamine emulation in ANNs for learning efficiency** - Emulation of the potentiating mechanism of dopamine in the connections of artificial neurons. The main ambition of this topic was to explore how the performance of conventional ANNs would be affected if a neuromodulator mediating connectivity were to be integrated into the system. Respectively, a dopaminergic trace was implemented in certain neurons throughout the network, forming "dopaminergic" variations of fully connected layers. This trace was used to manipulate the strengths of neurons' corresponding connections to the rest of the network. Additionally, the emulation of D1 and D2 dopaminergic receptor behaviors was also done in an attempt to improve the previous approach.

- **Epistemic emotion as a learning mediator for DL** - Integration of personal emotion scores of surprise and pride, factored based on the performance of a model to mediate some aspect of its learning. The goal of this approach was to provide insight into possible ways the learning autonomy of ANNs can be improved, and consequently the autonomy and adaptability of AI. In particular, a combination of RL and SL models as a mechanism capable of regulating its rate of novel training data was introduced. Here, the amount of novel data to be fed into a recognition network during its training was regulated by a Deep Deterministic Policy Gradient (DDPG) model based on its own epistemic emotions of surprise and pride. As a benefit, the DDPG model is agnostic to the architecture of the network being regulated, making it applicable to various other scenarios.

- **Artificial dreaming algorithm to prevent overfitting in ANNs** - Development of an algorithm for autonomous data augmentation by ANNs through the forcing of patterns on instances, based on the maximization of layer activations through gradient ascent, and posterior interpretation according to the network's current knowledge level. This process was designed for ANNs to perform sporadically if overfitting were to occur and attempt to reduce its effect, showing similarities with biological REM-phase dreaming in that the brain also undergoes these latency periods where random or warped data is analyzed for potential overfitting prevention. By integrating the augmented data in the training set the severity of overfitting can potentially be reduced. Additionally, the algorithm makes no assumptions regarding network structure or tasks being learned. Thus theoretically, it can be employed in virtually any feedforward architecture and corresponding problem with few tweaks.

1.4 Publications and Technical Contributions

1.4.1 Key Publications

The following list cites first author works and other collaborations that were published during the duration of the Ph.D., which are considered relevant to this thesis:

Journals

- **G. Assunção**, P. Menezes, F. Perdigão. "Speaker Awareness for Speech Emotion Recognition." International Journal of Online and Biomedical Engineering (iJOE), vol. 16, no. 4, pp. 15–22, 2020, doi:10.3991/ijoe.v16i04.11870.

- **G. Assunção**, N. Gonçalves, P. Menezes. "Bio-Inspired Modality Fusion for Active Speaker Detection," in Applied Sciences, vol. 11, no. 8, pp. 3397, 2021, doi:10.3390/app11083397.

- **G. Assunção**, B. Patrão, M. Castelo-Branco and P. Menezes, "An Overview of Emotion in Artificial Intelligence," in IEEE Transactions on Artificial Intelligence, vol. 3, no. 6, pp. 867-886, Dec. 2022, doi: 10.1109/TAI.2022.3159614.

Conference Proceedings

- **G. Assunção**, P. Menezes. "Intermediary fuzzification in speech emotion recognition," 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2020, pp. 1-6, doi: 10.1109/FUZZ48607.2020.9177699.

- **G. Assunção**, M. Castelo-Branco, P. Menezes. "ANNs Dream of Augmented Sheep: An Artificial Dreaming Algorithm." In Proceedings of the 2nd International Conference on Image Processing and Vision Engineering (IMPROVE 2022), pp. 135-141, 2022, ISBN 978-989-758-563-0, ISSN 2795-4943, doi: 10.5220/0011055700003209.

- A. Sorrentino, **G. Assunção**, F. Cavallo, L. Fiorini, P. Menezes. "A Reinforcement Learning Framework to Foster Affective Empathy in Social Robots." In: Social Robotics. Springer Nature Switzerland, 2022, pp. 522–533. doi: 10.1007/978-3-031-24667-8_46.

- **G. Assunção**, M. Castelo-Branco, P. Menezes. "Leveraging emotion-mediated exploration to adapt agent behavior." 2023 6th Experiment International Conference (exp.at'23), Evora, Portugal, 2023, *In Press.*

- **G. Assunção**, A. Sorrentino, J. Dias, M. Castelo-Branco, P. Menezes, F. Cavallo. "Adapting Behavior and Persistence via Reinforcement and Self-Emotion Mediated Exploration in a Social Robot." In: 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). IEEE, Aug. 2023. doi: 10.1109/ro-man57019.2023.10309410.

Preprints

- **G. Assunção**, M. Castelo-Branco, P. Menezes. "Self-mediated exploration in artificial intelligence inspired by cognitive psychology." arXiv preprint 2302.06615, 2023.

Workshops

- **G. Assunção**, B. Patrão, N. Gonçalves, M. Castelo-Branco, P. Menezes. "Sound-based Emotional Regulation for Improved HRI," in Workshop: Sound in Human-Robot Interaction, 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21), 2021.

- A. Sorrentino, **G. Assunção**, F. Cavallo, L. Fiorini, P. Menezes. "Modeling affective empathy by teaching emotion expressions to a social robot," in Workshop: Social Robots for Personalized, Continuous and Adaptive Assistance (ALTRUIST 2021), 13th International Conference on Social Robotics (ICSR), 2021.

1.4.2 Other Publications

The following list cites other works also published during the duration of the PhD program, but which are not directly within the scope of this thesis:

Journals

- **G. Assunção**, B. Patrão, P. Menezes. "Crowd Interest Mapping to Assess Engagement." International Journal of Online and Biomedical Engineering (iJOE), vol. 18, no. 2, pp. 167-180, 2022, doi:10.3991/ijoe.v18i02.25445.

Conference Proceedings

- B. Ferreira, **G. Assunção**, P. Menezes. "MIST: A Multi-sensory Immersive Stimulation Therapy Sandbox Room," in Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (GRAPP), vol. 1, pp. 160-168 ISBN 978-989-758-402-2, 2020, doi:10.5220/0009118401600168.

1.4.3 Other Contributions

Here are detailed other contributions besides publications, which were accomplished during this program:

- ROS node for wakeword detection and rotation towards its direction of arrival, based on Pocketsphinx [12] and time cross correlation between mic pairs on a MATRIX Creator [13] microphone array.

- Active speaker detection DL tool, combining facial and speech recognition over time with modality correlation via a novel fusion technique, as presented in [14].
- Automated data mining technique based on the combination of metaheuristic techniques such as Simulated Annealing or genetic algorithms, and a conventional ANN.

1.5 Thesis Outline

Here is provided a summary general structuring of each topic presented in this document, in addition to related publications and contributions. The outline is as follows:

- **Chapter 2 - DL Basics and Functioning** This technical section introduces the DL topics employed in the works that constitute this thesis, concisely and objectively.

Part I - Biological Overview

- **Chapter 3 - Neurophysiology** This summary overviews concepts essential to understanding biological neural circuitry, particularly involved in emotional processing and decision-making, as well as how learning occurs and how real brains adapt to deal with real-world variability.
- **Chapter 4 - Psychology** This section provides insight into the interpretative take of Psychology over emotion, reviewing the cognitive-motivational-emotive system proposal as well as the effects learning has on the body and behavioral response.

Part II - Emotional Competence

- **Chapter 5 - Fuzziness of Emotion** This chapter presents an approach to speech emotion recognition, combining deep learning and fuzzy c-means clustering to account for interstate information when assessing emotional utterances.
- **Chapter 6 - Empathy in Social HRI** This chapter proposed a reinforcement learning technique for social robots to develop empathy, via direct human feedback received when attempting to match facial expressions with their respective users in an interactive scenario.

Part III - Adaptability

- **Chapter 7 - Emulating Dopamine** This chapter describes an approach to mimicking dopamine effects in artificial neural networks, intending to improve learning efficiency. A basic dopaminergic trace is first developed, followed by integration of D1 and D2 receptor emulation.

- **Chapter 8 - Exploration from Internal Drives** This chapter reviews a deep learning technique, combining insight from cognitive psychology and neurophysiology, and developed for artificial agents to learn useful correlations between artificial emotion and exploratory behavior, via reinforcement.

Part IV - Final Remarks

- **Chapter 9 - Artificial Dreaming** This chapter describes an autonomous data augmentation algorithm designed for artificial neural networks to benefit from the overfitting prevention effect that real dreams are theorized to induce on a biological brain.
- **Chapter 10 - Final Remarks** This section brings the manuscript to a close by summarizing the achievements of the work presented, as well as providing paths for future research to continue development.

Chapter 2

Deep Learning Basics and Functioning

Contents

Neurophysiologic concepts were first adapted for computational simulation in the 1940s and 50s, with researchers striving to better understand basic neuro-functioning. Yet even after the formulation of the artificial neuron (perceptron) [15], ANN development did not gather much support due to the technological setbacks of the time. It was not until the late 20^{th} and early 21^{st} century that DL, a field focusing on optimization via multi-layered ANN architectures with representation learning, became a mainstream approach to engineering problems. This chapter presents some theoretical background of DL as well as the techniques that were employed for this thesis.

2.1 Representation learning

This concept is a cornerstone of ML and more importantly DL. It is characterized by a form of learning based on data abstractions from which useful information can be easily extracted and serve as suitable inputs for predictor/decision-making models [16]. Naturally, it is very typical of stochastic models, where representations can capture a posterior distribution that fits input data instances to a good extent, as well as suggesting priors that may not be data specific but are still useful to learning related tasks. Hence, with neural networks, data representations become increasingly refined as they progress through the models and as training optimizes layer weights. Hence, deeper or wider architectures may generate progressively more abstract features, even from complex data types whose description is not straightforward. Examples of this include supervised learning, where data representations commonly serve as descriptors of meaningful cues to the learned task, and reinforcement learning, where internal representations characterize the intrinsic or extrinsic value of a state or chosen action.

Moreover, ANNs also have the ability for representation reuse or distributed use, which allows different perspectives over the same data description and can potentially lead to better results. Nevertheless, the relationship between representation abstraction and suitability for task learning is not linear, and thus care must be taken when designing neural networks for specific tasks. In terms of emotion, a naturally subjective concept, representation learning appears to be a suitable tool as it can capture the abstractness of this concept.

2.2 ANNs & Supervised Learning

The SL class of ML encompasses techniques to model the relationship between data instances and their labels (the training set), for subsequent usage as predictors for unlabelled instances of the same or similar type (the testing set). The data-label tuples are considered to be sampled from an independent and identical distribution (*i.i.d.*) which remains the same over time, so the split between and generalization from training to testing instances can be considered valid. When progressing to DL, with basic feedforward ANNs, layers are composed of basic units called neurons. The output of these is typically a linear combination of its inputs rated by the intensity of the connections (weights) with the respective previous neurons and activated by a function which determines if/how the information will progress to the following neurons. This computation, resembling the functioning of a real neuron, is repeated along the network structure until a final output is obtained. A chosen loss function L is then used to quantify the deviation of the outcome (prediction error) produced by the network and the expected value which is the corresponding label. The gradients of L with respect to a layer's weights can then be used to update them towards optimized values, with the chain rule being sequentially used as we move upstream until all weighted layers have been updated. This is done iteratively until the network reaches an acceptable performance for the task it is being trained for. Selecting an appropriate loss function depends on the nature of the task, as optimization is performed differently for distinct types of problems (e.g. regression *vs* classification).

ANN architectures have varied considerably, with modifications presented to match demands found for specific problems. For instance, in Convolutional Neural Networks (CNNs) a layer's neurons constitute filters that are convolved with the input data to reduce complex patterns into smaller simpler patterns. With training, filter weights gradually begin to accentuate the patterns they generally detect and become highly useful for image analysis. In residual networks (ResNets), neurons may connect with others not necessarily in the following layer, via *skip* connections. This methodology allows for deeper and more robust networks since fewer connections equate to less feature space exploration and gradients may flow upstream through the *skips* rather than vanish entirely. Recurrent Neural Networks (RNNs) relax the independence between input and output since neurons consider feedback from prior samples as additional input to the analysis of a current sample. This enables optimization according to the temporal correlations of data samples and can be useful for the network to recognize any sequential characteristics of that data. In addition to these network types, several other examples exist, being either sub-variations, a combination of one or more types, or a distinct architecture altogether. However, addressing them would be out of scope for this thesis.

2.3 Reinforcement Learning

RL, as described by Sutton and Barto [17], is arguably the branch of ML which most closely resembles the neurological learning processes described in Chapter 3. In broad terms, this methodology usually considers an environment on which an agent, with no prior knowledge, performs an action $a_t \in A$ leading it from one state s_t to the next, in a finite or infinite state space S. The outcome of an action is signaled by a reward function $r(s_t, a_t)$, and thus the return or value from a state can be regarded as a sum of discounted future rewards with a discounting factor $\lambda \in [0, 1]$:

$$R_t = \sum_{i=t}^{T} = \lambda^{(i-t)} \cdot r(s_i, a_i) \tag{2.1}$$

With this knowledge, reward can be interpreted as signaling what is immediately good and coveted by an agent. Separately, value refers to potential future rewards associated with a state, and thus a value function is meant to determine the worth of a state.

The *i.i.d.* perspective does not apply to problems in RL. Instead, these are commonly devised in the form of a Markov Decision Process (MDP) where given an initial state distribution $p(s)$, $p(s_{t+1}|s_t, a_t)$ represents the probability of the agent transitioning to state s_{t+1}. The behavior of the agent at a reached state, or mapping from that state to the probability of an action, is commonly defined as a policy $\pi : S \to \mathcal{P}(A)$. Moreover, by designing a signal r consistent with the objectives of a task, it becomes possible for agents to approximate an optimal policy for it, which maximizes the expected return $J = \mathbb{E}[R_1]$ from the initial distribution. This represents the major objective of the RL paradigm and is affected by an exploration-exploitation dichotomy. At earlier stages of training, exploration is usually preferred but should fade as the agent gathers knowledge from the environment and later relies more on exploitation. Without this exploratory preference, an agent is characterized as greedy as it relies (almost) exclusively on exploitation. Regardless, by the end of training the agent should be able to predict the value associated with a state-action pair and appropriately decide on an action for each state. This process mirrors that of associative learning via primary and secondary reinforcers in a real brain, as an agent appraises state-action pairs and employs this knowledge in the decision-making process so that it can meet its goal and/or maximize its chances of survival. With this premise, it seems plausible that RL methodology would be an adequate framework for the emulation of biological processes in AI research.

Methodologies can be categorized as on-policy or off-policy. The former refers to techniques that actively and directly improve upon the probability distribution π which the agent uses for decision-making, as training progresses. On the contrary, off-policy techniques attempt to improve performance by updating a target policy distinct from the behavior policy π, thus being independent of the agent's action. There is a range of advantages and disadvantages to both types of RL methodology.

For instance, off-policy methods may inaccurately estimate the value of an action if the behavior and target policies differ a great deal. This does not happen with on-policy techniques since there is a single policy to use. On the other hand, on-policy techniques may become trapped in local minima, as the policy used for action selection iteratively becomes better at reaching a certain result. This drawback is prevented in off-policy methods by greater flexibility and exploration, as learning assumes the use of a greedy (target) policy while action selection employs another (behavior) policy.

In addition to policy, RL methodology may be further characterized by the availability of a model, which the agent can use for predicting environment responses. This is done via the future reward associated with an action selected for some state. In model-based approaches, algorithms either employ a model of the environment available from the get-go or one may be built based on observations performed as the agent explores the environment. On the contrary, model-free algorithms rely solely on sampling the real environment and never reference a model to predict the next state/reward. Considering the nature of most real-world scenarios and the inaccessibility or complexity of generating a model describing them, techniques in this area are largely model-free. They are also categorized as deep RL, as deep neural networks often compose the models employed. Further terminology includes history, referring to the sequence of observable variables (e.g. action, reward) up to a certain moment, replay, for when an agent reuses history information to learn rather than the most recent sample, and others specific to the RL techniques they describe, such as Deep Q-Learning Networks (DQNs) or DDPGs.

2.3.1 Deep Q-Learning Networks

Q-Learning [18] was first introduced as a model-free, off-policy algorithm for learning a function $Q(s_t, a_t)$ which models the value associated with some state-action pair in discrete space, with which the best action is chosen for the current state. This decision depends on a look-up table, where the reward values for each (s_t, a_t) pair are stored. The process aims to maximize reward and optimize this function by iteratively updating it according to knowledge gathered from the environment in the form of temporal difference (TD) of optimal and Q-values experienced for actions based on a deterministic policy, using:

$$Q^{new}(s_t, a_t) = (1 - \alpha) \cdot Q^{old}(s_t, a_t) + \alpha \cdot \Big[r(s_t, a_t) + \gamma \cdot \underset{a'}{max} Q(s_{t+1}, a') \Big] \tag{2.2}$$

A learning rate α determines the weight of the update over Q when a_t is chosen at state s_t, preventing the algorithm from racing to a poor solution. The update is based on the accumulation of the obtained reward and the estimate of the optimal future value at the next state assuming a greedy policy. In other words, this second term considers the Q-value associated with the best possible action at state s_{t+1}, so the agent may consider possible future rewards before performing a step. A discount factor γ strains the effect of this consideration on the immediate decision-making of the agent. This is so that earlier rewards the agent discovers may cascade down to the current state and still be valued. An issue occurs with this process when an agent becomes greedy and stuck selecting the action corresponding to the maximum Q-value at any given state, even though these values are still not optimal. To counter this, a decaying factor ϵ may be used to determine whether or not the action chosen should be randomized or according to policy. Thus, exploration is initially boosted and scales down over time so the agent may focus on exploitation of accumulated knowledge at later steps.

When dealing with high-dimensional problems, it becomes infeasible to approximate a Q-function using information from look-up tables. As such, DQNs were introduced as a combination of Q-learning and deep ANNs to approximate the Q-value function for a scenario. Input consists of the state, while the network outputs the Q-values for each possible action, so the agent may then decide what is the best course of action. As the training of this network progresses, output values should be near those produced by (2.2), and therefore the loss function should be:

$$L(\theta) = \mathbb{E}\left[\left(r(s_t, a_t) + \gamma \cdot \underset{a'}{max} Q(s_{t+1}, a'|\theta) - Q(s_t, a_t|\theta)\right)^2\right] \tag{2.3}$$

The first two terms constitute the target value, whilst the third term is obtained directly from the network as a prediction. Therefore, the subtraction enables minimization of the loss and eventual convergence to a solution. While guaranteeing performance here is impossible with non-linear and large function approximators such as ANNs, convergence is possible as several applications have shown to this day. Otherwise, techniques such as replay buffers and target networks may be used [19] to stabilize learning.

2.3.2 Actor-Critic Methods

Another methodology in deep RL consists of two separate structures working in tandem and which serve as the policy and estimate of value functions, respectively called actor and critic, in a dynamic resembling that of the basal ganglia (BG) process described in section 3.3. While both the actor and critic receive the state s_t as input, the former decides on an action a_t which the latter then critiques by outputting the associated value $Q(s_t, a_t)$. This allows the critic to determine whether the chosen action led the agent to a better or worse situation, and pass on the TD error to the actor. Given how this occurs at each timestep, actor-critic methods may be applied to continuous action spaces, unlike Q-learning. The update over critic parameters employs the Bellman equation, while the update over actor parameters (i.e. the probability distribution of actions) is based on the policy gradient theorem [20], and aims for actions with a higher expected reward at a state to have a higher probability value. The dynamic of this methodology is exemplified in Fig. 2.1. Learning in natural actor-critic methods is considered on-policy, given how the critic learns based on the same policy used for action selection - the actor. Still, variations have been proposed, including off-policy versions [21], which have the advantages against on-policy counterparts mentioned previously while also being robust to noise in updates. Evidently, in off-policy versions, the critic learns a value estimate for a policy distinct from the behavior policy, which is updated by the actor based on said estimate.

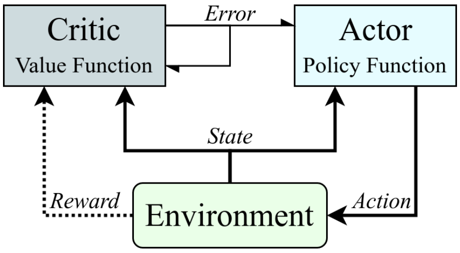


Figure 2.1: Simplification of the actor-critic architecture and relationship with the environment.

2.3.2.1 Deep Deterministic Policy Gradients

First proposed in [22], DDPGs are a variation of policy gradients employed in actor-critic, with a valuable level of exploration stemming from a stochastic behavior policy. However, Bellman-based critic updates consider a deterministic target policy estimate, which is far easier to learn. Assuming a neural network scheme, with one parameterized by ω as the critic model Q_ω, and another by θ as the actor model μ_θ (deterministic policy mapping state s_t to action a), the Deterministic Policy Gradient (DPG) is obtained from applying the chain rule to a performance objective, here being the expected return $J(\mu_\theta)$:

$$\nabla_\theta J = \mathbb{E}\left[\nabla_\theta \mu_\theta(s_t) \nabla_a Q(s_t, a|\omega)|_{a=\mu_\theta(s_t)}\right] \tag{2.4}$$

Given how this gradient describes the performance of the policy, it is used for updating the actor parameters. In addition, the issue with using large function approximators such as ANNs still holds for DDPGs as convergence cannot be guaranteed. As such, the proposed architecture for this technique also uses a replay buffer to reduce the variance from temporal correlations. In it, a $(s_t, a_t, r(s_t, a_t), s_{t+1})$ tuple is recorded with each step. Subsequently, a N-long mini-batch of index i may be obtained from the buffer to update network parameters. Besides the replay buffer, target networks are also used to regularize learning. These networks, respectively $\mu'(s_t)$ and $Q'(s_t, a)$, copy the weights of their actor and critic counterparts and are used to compute the TD target by summing their outputs with the reward for each sample. The loss of the critic model is then based on this target and the output $Q(s_i, a_i)$ obtained for the i^{th} sample. Thus, for each mini-batch:

$$L = \frac{1}{N}\sum_i \left(\underbrace{r_i + \gamma \cdot Q'(s_{i+1}, \mu'(s_{i+1}|\theta')|\omega')}_{\text{target}} - Q(s_i, a_i|\omega)^2 \right) \tag{2.5}$$

Gradients from this loss function are then used to update critic network parameters. Subsequently, at the end of a training step or another pre-defined interval, the target networks are updated with the newly calculated actor and critic weights, respectively. However, these updates are softened by a factor $\tau << 1$ so learning is slower and consequently more stable. Finally, a noise process may be integrated with the behavioral policy of the agent to boost exploration, taking advantage of the off-policy nature of DDPGs.

Part I

Biological Overview

Prologue

AI and robotics have gained immense traction over the past decade, producing increasingly successful applications as we strive to explore and exploit various new possibilities. In this sense, realistic behavioral attributes and learning autonomy are natural next steps for their development. Yet, in terms of neural adaptability and emotion, failing to understand the link between Psychological subjective influence and Neurophysiologic objective processes is a potential reason why research into the artificial implementation of such traits is only now emerging. Plus, often publications in areas of biological study do not target interdisciplinary interest. Consequentially, its intelligibility may require additional time and effort before it may be useful as a basis for AI development. To mitigate this and other issues limiting interdisciplinary synergy, this part of the thesis introduces a general overview of neurophysiologic and psychological concepts relevant to the presented DL approaches, which may also serve as a starting point for additional works in bio-inspired AI. Accordingly, no assumptions are made whatsoever regarding familiarity with terminology or knowledge of either area.

The information here provided is adapted from surveyance work, partially published already in [23], but ongoing. It benefits from the complementarity of both Neurophysiology and Psychology, as in the field of Cognitive Neuroscience. The first chapter scrutinizes the underlying biology of the brain, focusing on learning and emotion, and how these processes affect and lead the body to act. On the other hand, the second chapter studies the causes and consequences of said processes in human life, attempting to model them. These two information groups are useful as they demonstrate not only what could be attained in artificial agents emulating biological characteristics, but also how those may be exploited in research to fulfill other objectives.

Chapter 3

Neurophysiology

Contents

One way to further autonomous learning and behavior in AI and robotics is to look for solutions where these traits are already established. The brain already boasts unparalleled autonomy and learning capabilities. Thus, this chapter explores neural circuitry and processes related to adaptability and emotion, introducing key parts of their biological foundation as a basis for AI emulation. Naturally, the overview is not exhaustive and instead is meant to highlight certain topics of emotional neural processing, neural plasticity and modulation, and reinforcement learning in the brain which are considered relevant for the DL work developed in this thesis.

3.1 Basic Functioning

The brain is essentially a large compound of basic cell units, known as neurons, occasionally compartmentalized into specialized and interconnected structures, which form highly convoluted neural networks. The neuron is commonly modeled as an accumulate-and-fire trigger cell [24], receiving input from either other neurons or receptor nerve cells. In a process called a chemical synapse, information is conveyed via chemical messengers, the neurotransmitters, which traverse thin clefts in between neurons and reach receptors in cell body ramifications called dendrites capable of converting those chemical signals into small electric impulses. This process either depolarizes or hyperpolarizes the postsynaptic cell, respectively representing either inhibitory or excitatory transmission of information. Should it lead to a strong enough electric disturbance, the resulting action potential (the *spike*) is conducted to a set of terminals where it initiates the release of the mentioned neurotransmitters via gated channels.

The described synaptic process varies considerably in terms of speed, bursting activity, structural efficiency, and other factors. These fluctuations are mediated by a set of chemicals known as neuromodulators, analogous to neurotransmitters. Unlike these, however, neuromodulators are diffused, not necessarily at synaptic sites, and received by a distinct type of receptors. Consequently, they may affect a group of adjacent neurons or even be widespread enough to reach proportionally distant neural structures. Moreover, effects over synaptic or even cellular properties tend to be long-lasting, causing a potentiation or attenuation of neuron connections. This process, often referred to as synaptic plasticity, depends on factors such as membrane excitability, synaptic transmission, and integration, the sensitivity of receptors to neurotransmitters as well as the probability of neurotransmitter release, and several others. Dopamine is an example of a neuromodulating chemical, having a key role in motivation, reinforcement, and reward but also mediating the plasticity of neural connections. As such, it is addressed below to better understand the methodology implemented for this thesis.

This summary of a brain's functioning is enough to understand how its sections specialized and how certain elements came to be, depending on the information flowing through and which body parts they interface with, to allow beings a better chance of survival in the real world. Emotion is an example of these elements, strongly modulated by specific positive or negative instrumental reinforcers [25]. It is processed in regions commonly referred to as limbic structures and it significantly impacts learning and behavioral aspects relevant to this work, thus also being shortly overviewed hereafter. Regarding the neural information presented below, it should be noted circuitry and operation expositions are not fully comprehensive and reflect only aspects of functioning pertinent to the work developed in this thesis. Further analysis would be out-of-scope for this document.

3.2 Neural (Limbic) Circuitry and Emotion

The historical idea of a well-bound limbic system has been relaxing over the years [26, 25, 27]. Thus, when it comes to emotion, processing is generally considered to be decentralized and following a tiered fashion. First, sensory cortices receive a stimulus from their corresponding organs, identifying it. Following this, the information is then assigned value by limbic structures which largely account for the appraisal, storage, and recollection of environmental stimuli representations refined for emotionally motivated learning and behavior. Primarily, this encompasses the amygdala, nucleus accumbens (NAc), and the medial prefrontal/orbitofrontal regions of the prefrontal cortex (PFC), but also other structures such as the ventral tegmental area (VTA). Finally, decisions are made as to how the body should react or behave, in response to the perceived stimulus. The fallout of body action is likewise captured by sensory cortices [28] and then used to feedforward update those same neural structures. A diagram of this emotional system, along with relevant inter-structure connectivity and approximate locations within the brain is shown in Fig. 3.1.

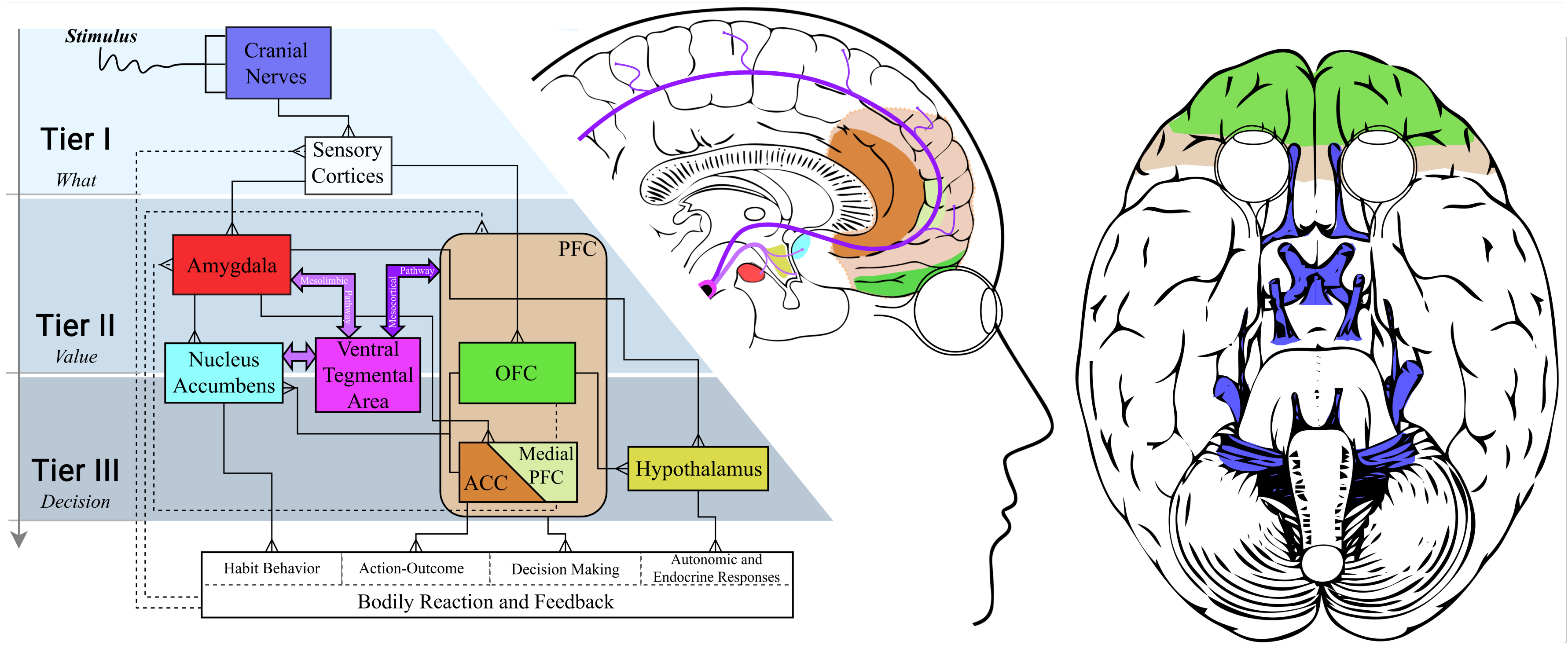


Figure 3.1: Tiered schematic of interaction between brain systems intervening in emotion processing on the left, depicting information flow from stimulus capture to bodily response. Color coding of each structure block is used to represent the approximate location on the lateral section and ventral views of the human brain on the right. The full array of connections between the represented structures is not provided to avoid cluttering and focus attention on the analogy with artificial intelligence. Dashed lines represent feedback to upstream structures. As presented in [23], based on [29], [30] and [25].

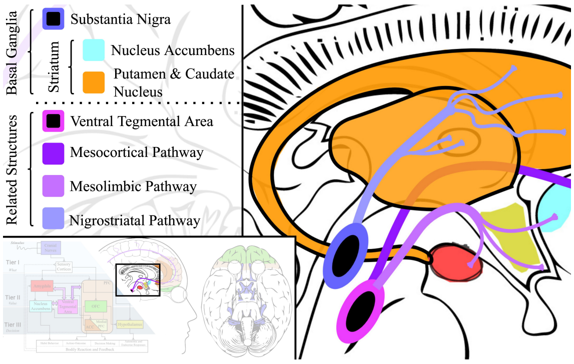
Among the structures with the highest intervention rate in emotional processing, the amygdala is a major processor of conditioned reinforcers that performs an appraisal of stimuli relevance regarding appropriate body action [31]. Moreover, sections of the amygdala (e.g. the basolateral region) maintain the reinforcing properties of a stimulus as part of their linkage to sensory structures elsewhere in the brain [32, 33], forming a type of affective memory. Consequently, the amygdala enables appetitive/aversive discrimination of stimuli [34] and attentional capture through prolongation of response times [35]. This causes emotional salience (i.e. filtering) of the environment [30] and contributes to reactive behavior, impacting affective and low-level aspects of social conduct [36], as well as mediating associative learning [37]. In terms of connectivity, the central nucleus of the amygdala mediates arousal and autonomic response systems such as the hypothalamus [32], whilst projection to higher-level neural structures such as the medial prefrontal cortex (mPFC) relates more with cognition and knowledge consolidation [38]. Given this duality, the decision tier encompasses a more intuitive side where the autonomic component enables us to act based on what is colloquially known as *gut feeling*, which contrasts with a cognitive control section for rational thinking. Both sides influence affective valuation as it depends on robust feedback connectivity from structures at this level.

Likewise to the amygdala, the PFC also boasts major emotional processing through its anterior cingulate cortex (ACC) and orbitofrontal cortex (OFC) areas [32]. Unlike the amygdala, however, the OFC is also involved in affective storage and reinforcement [25, 39] yet by stocking representations of value associated with reward/punishment outcome. In addition to influencing autonomic and motor responses [37] analogously to the amygdala, the OFC also enables high-level perceptual integration and cognitive control. Moreover, firing from both these structures activates the ACC and its adjoining regions, which compute action costs and enable appropriate goal-directed behavior [40]. Ergo, the ACC facilitates learning of relevant action-outcome patterns by actively monitoring associated error [39]. Not only does this process affect mood but, based on intrinsic value representations, it allows us to decide on what is most advantageous from conflicting environmental cues and plan for survival [32, 33, 30].

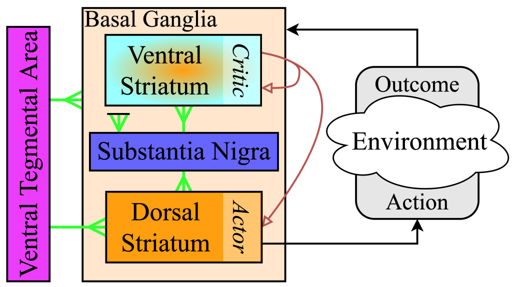
The NAc section of the Ventral Striatum (VS) is another meaningful region for emotional processing, receiving appraisal patterns from the amygdala and PFC. With these, it codifies the incentive value of emotionally significant stimuli [37]. This feature enables the accommodation of reward delays, which helps refine preparative action as well as switch between goal-oriented and habit-based behavior [32]. Similarly, the VTA also affects this cognitive-intuitive behavioral switch, via the mesolimbic and mesocortical pathways. This is because the former affects limbic circuitry, such as the amygdala and NAc [41], while the latter projects mainly to the PFC. The amount of dopaminergic projections accommodated by these pathways is rather high, also affecting decision-making factors [42], such as arousal, reward, and Prediction Error (PE), and mediating incentive salience.

3.3 Dopamine and Related Structures

As introduced above, dopamine as a neuromodulator is a major enforcer of neural reconfiguration through modulation of synaptic plasticity. Within the brain, it is mainly sourced from cell groups dubbed dopaminergic, such as the VTA, a small-scale structure named substantia nigra (SN), and the NAc in the VS [43]. Both the former, through the mentioned mesolimbic and mesocortical pathways, and the latter, via the nigrostriatal pathway, are responsible for a large majority of the dopaminergic projections occurring in the brain. An adjusted detail of Fig. 3.1 showing these structures, in particular, is provided in Fig. 3.2A.



(A) Approximate locations



(B) Actor-Critic System

Figure 3.2: Overview of the dopamine system and related actor-critic modeling of the basal ganglia. (A) shows the approximate location of structures most responsible for dopaminergic projection in the brain, and corresponding connections to BG components. Based on [25] and [43]. (B) demonstrates a simplified system where a BG can learn an appropriate action, encouraged by dopaminergic projections (green) which vary depending on the corresponding outcome's PE (red). Based on [44] and [45].

At a cellular level, dopamine signaling presides over a synergy between two families of receptors, $D1$ and $D2$, respectively with potentiating or attenuating capabilities [46]. Succinctly, when activated, $D2$ receptors at postsynaptic cells mediate inhibitory transmission from presynaptic neurons, while on the contrary $D1$ receptors in excitatory synapses extend a transmission further than it would otherwise last. Moreover, dopamine concentration strongly determines the activation of these receptors [47], with $D1$ activation occurring at low dopamine levels while higher concentrations target both $D1$ and $D2$ receptors, albeit with a premium over the latter family. This behavior of $D1/D2$ receptors is fundamental for Spike Timing Dependent Plasticity (STDP) to occur, a process wherein synaptic strength is adjusted based on the relative timing of neuron spikes [48]. Not only does STDP make coincident stimuli increasingly associable [46], demonstrating dopamine's influence over several components of reinforcement learning, it also represents a cornerstone of Hebbian learning, the theory that continued activation of pre and postsynaptic neurons potentiate the synapse in-between them [49]. Thus it also exposes the strong influence dopamine has over the structural tuning of a network.

At a structural level, dopaminergic projections from the VTA and SN heavily target components of the BG. This group of nuclei influences motivation and decision-making, being generally posited as modulatory in the action selection process of the brain [44]. Further, BG functioning has been modeled as an actor-critic RL process where an action's predicted and real outcomes are compared and learning occurs by trial and error. Here the Dorsal Striatum (DS) typically represents the actor, learning stimulus-action pairs and regulating motor function. This portion works in parallel with the VS as a critic, itself composed of parts such as the NAc which codifies the action's associated value based on received appraisal patterns as explained previously. Both striatum sections have been observed receiving signaling from dopaminergic neurons in the SN and VTA, proportionately to the reward PE associated with a chosen action [45]. This transient dopamine influx or lack thereof, associated with $D1/D2$ receptor behavior, regulates the expedition or gating of that same action by the actor-critic pair in future similar scenarios. An exemplification of this process is modeled in Fig. 3.2B for easier understanding.

3.4 Learning through Reinforcement

Learning here refers to how knowledge is obtained and solidified in the brain through an amalgam of neural processes involving reinforcement, reward-based, and associative algorithms [32, 50]. These learning mechanics particularly are highly subject to emotional influence and most often take place over the neural circuitry described above.

Stimuli captured from the environment inadvertently triggers certain involuntary responses (e.g. the sense of taste) and a consequential feeling of pleasantness or lack thereof, in a process known as primary reinforcement. In case secondary but simultaneous stimuli also occur (e.g. sight of food), the latter becomes associated with the emotional value the primary stimuli was appraised with, as well as the autonomic response it induced. The secondary stimuli are then dubbed secondary reinforcement [25, 39], as its influence over the body is based on its relation with primary reinforcers. The creation of these associations constitutes knowledge acquisition by the brain, in the form of an emotional dimension. This is updated continuously so that new appraisals (i.e. predictions of value) more closely fit the information stored about previously experienced outcomes. This knowledge is employed in decision-making so that a being's actions over its environment meet some goal but also maximize its chances of survival. The process of selecting an action is based on the anticipated reward, which itself is represented internally by the emotional value of the action's expected outcome [30]. Naturally, the more an environment is trialed, the better reward prediction becomes. This process is summarized in the diagram of Fig. 3.3.

This learning mechanism is made possible by the neurotransmission chemical process, described previously, which conveys information to receiving cells. Through it, neurons may signal different aspects of reinforcement and decision-making [51, 52], in addition to potentially inducing an emotional impact [53] across different brain regions. Dopamine, overviewed above in terms of its neuromodulatory properties, is also heavily involved in the broadcast of PE [50] as a neurotransmitter.

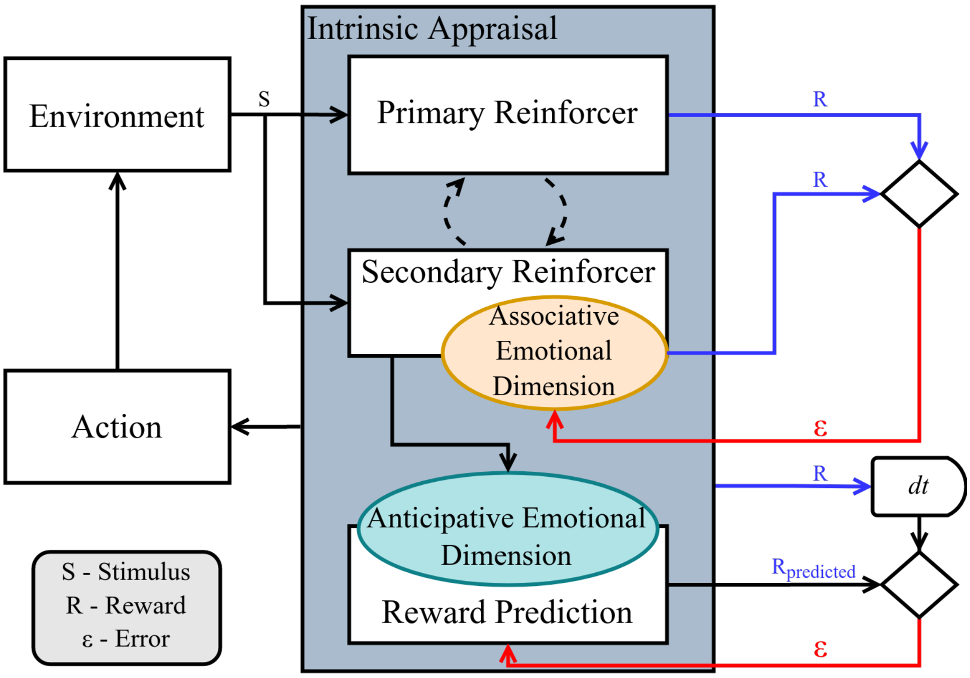


Figure 3.3: Learning model based on emotional appraisal and reinforcement. Associations between primary and secondary reinforcers may be created so that environmental stimuli may be emotionally appraised and appropriately responded to, based on the predicted reward of their corresponding outcome and avoiding involuntary reaction. As presented in [23], based on [25] and [50].

Progressive transfer from post-reward to pre-stimulus activation has been observed for dopamine [54], demonstrating its importance for communicating anticipation and forming predictions, respectively. Before dopaminergic error prediction, stimuli salience may be based on another neurotransmitter named norepinephrine. This is because its diffuse projections from the locus coeruleus structure of the brain boost attention and arousal. These elements are pivotal to attentional reward components [55], leading to the activation of dopaminergic neurons. Neurotransmitter roles on emotional salience, learning, and attention can be further explored in [56, 50], as a further review would be out of scope for this thesis.

Chapter 4
Psychology

Contents

Bio-inspired processes, given their function role in human cognitive processes, are naturally causal of behavioral tendencies, expression, and what constitutes the social construct we live in. Simultaneously, these factors determine the contextual information that the brain captures, triggering the processes described previously. However, Neuroscience alone may not provide sufficient attention to these properties and study them in correlation with neural inner workings. For example, when dealing with emotion as an entity, it becomes necessary to model it in computationally applicable terms, and not simply as a consequence or effect of relationships between limbic structures. Thus this chapter provides a psychological perspective on topics related to learning and emotion which are important for some DL areas. With equal reasoning to the previous chapter, this overview of psychological concepts addresses only what is considered relevant to the computational work developed in this thesis.

4.1 Basis Outline

Despite decades of focus, lack of consensus is prevalent and a direct consequence of the perplexing multitude of theories for emotion [57] proposed in Psychology. Yet it is clear how most research agrees on appraisal theory [58], where emotion directly or indirectly stems from the personal meaning an individual assigns to the context of some triggering event. For instance, Keltner and Gross' construct of emotion [59] is consensual, envisioning patterns of perception, communication, and action which are episodic and triggered by challenges/opportunities of a physical or social nature. Thus it makes sense that emotional states would be diverse. Moreover, emotion itself weighs in as feedback to the subjective experience or outcome of an event's appraisal. This process leads to physiological responses and consequential leakage of multi-modal cues, internal and external, as well as diverse behavioral tendencies [60, 61, 62, 63] which translate to action. Through it, emotion motivates behavior by coordinating body systems to react toward some situation in an attempt to optimize a person's chance of survival or goal achievement [64, 65]. This link between context and the intra-individual factors that motivate emotion response generation is better understood through Fig. 4.1, demonstrating the relationship between cognition, motivation, and emotion [66].

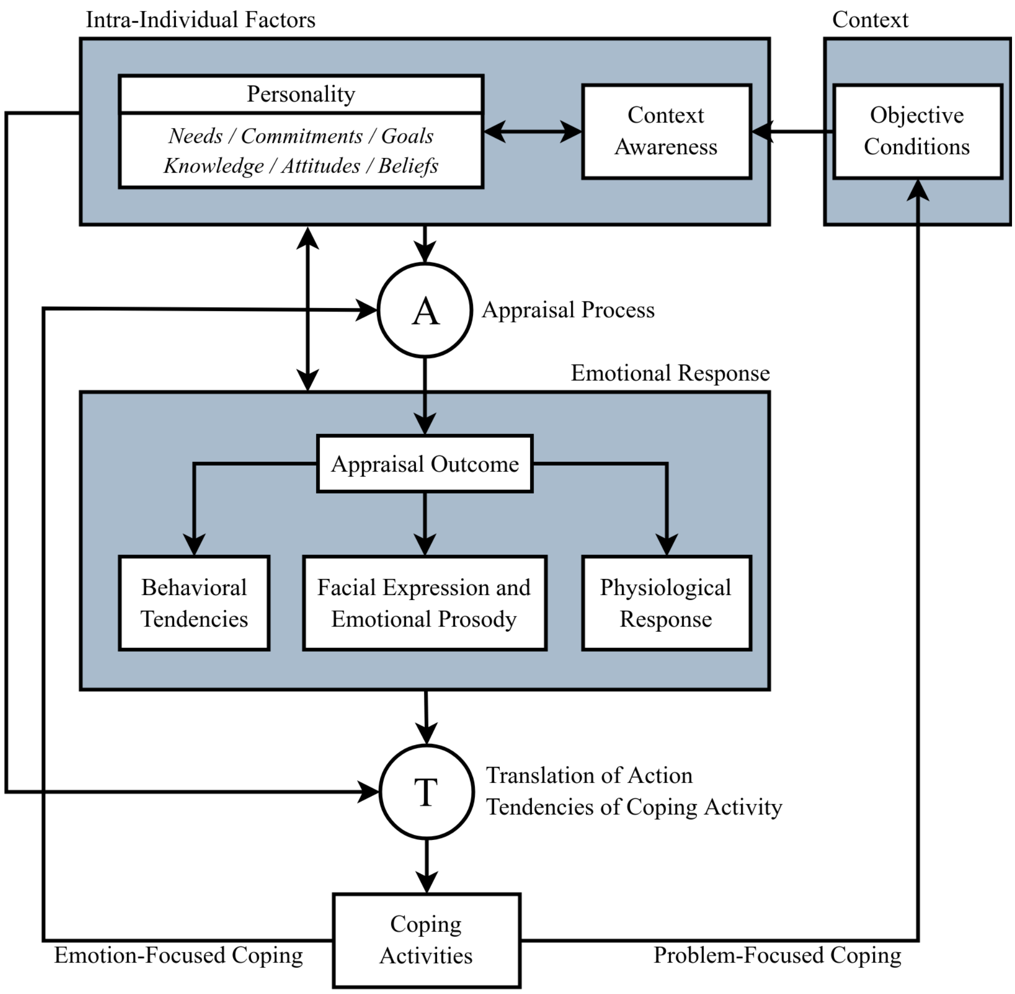


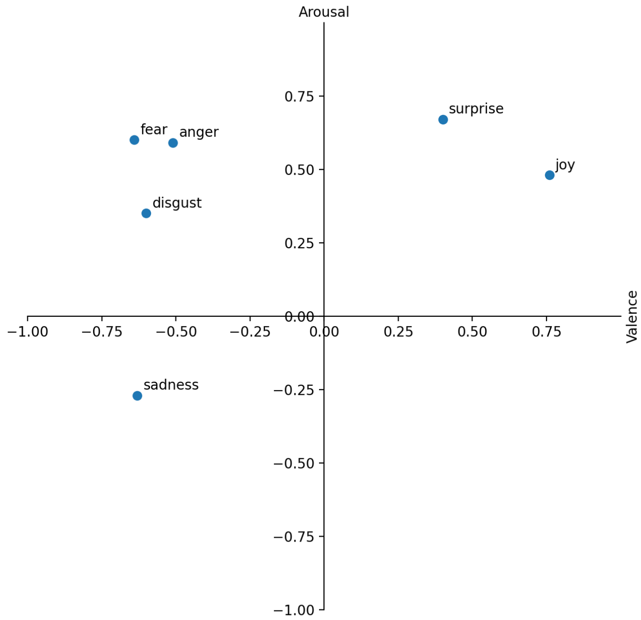
Figure 4.1: The Cognitive-Motivational-Emotive system, abridged from [66]. The emotional appraisal of some scenarios is directly affected by the context it occurs and the individual's personality. A relationship between these two, determining how the person perceives the context, is also influential. This process elicits a response from the body in the form of autonomic reactions such as facial expressions and physiological changes, in addition to motivating some specific behavioral tendencies. The combination of those components translates to an action meant to cope with the appraised situation in a manner again dependent on intra-individual factors. This in turn will affect the appraisal process of subsequent scenarios.

The observation of populational trends during physical/social manifestations is generally what enables emotional states to be understood and cataloged since the analysis of each expression separately usually causes a lack of consensus [67]. Additionally, emotions tend to be regarded as innate, with separate people manifesting them and the corresponding physiological reactions identically when experiencing similar situations. There is a degree of complexity associated with each emotional state. Some elicit a more overt response (e.g. anger) and can be considered primary, while other secondary ones require further analysis (e.g. boredom). In objective terms, the visual and auditory modalities arguably provide the most readily available information about emotional states, with analysis of physiological signals (e.g. galvanic skin response) following in a close second. For instance, behavioral responses such as body posture, facial expression, or prosody variation are commonly observed and may constitute input to a model attempting to understand and predict human emotional states.

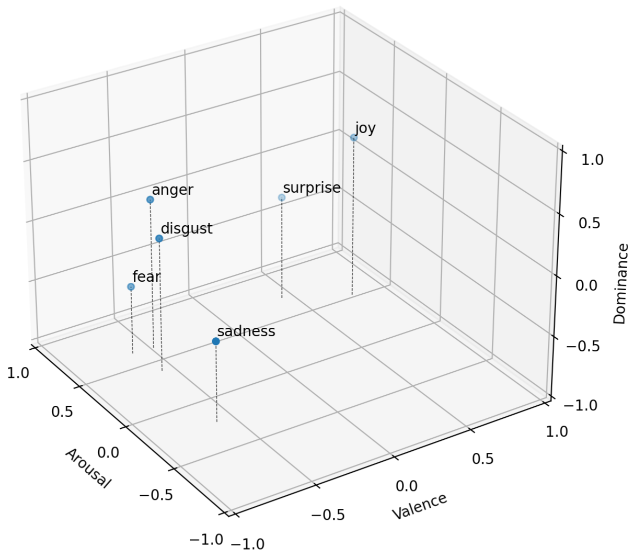
4.2 Modelling

Emotions can be more primitive and influence our automatic survival mechanism in threatening situations (e.g. fear), with heavy autonomic effects (e.g. heart rate, body temperature), intrinsic scenario appraisal, and adequate action selection known as the fight-or-flight-or-freeze response [68, 69]. These are typically well-known states, easily recognizable. Other states are of a social nature (e.g. shame) and develop as we fit the environmental settings and interactions we are exposed to during growth [70], with identification not being as straightforward. Nonetheless, there is a positive or negative valence to emotion, depending on the situations it correlates with, either successful or threatening respectively [65]. Likewise, its arousing effect on the body can be used to characterize the specific state. Moreover, physical or body-based metrics can be considered as emotional features and used to monitor the internal and external components of a state, as well as reporting intrinsic appraisals during particular events or experiences [71, 72].

In terms of objective modeling, emotion is generally divided into two major subgroups. The most well-known and historically adopted one corresponds to a discrete categorization of emotions, where these are commonly associated with a finite closed set of archetypal states (e.g. happy, sad, anger, surprise) defined as physiological-related and measurable by Plutchik [73] and Ekman [74]. Typical examples consider around six archetypal states as universal, though the set size varies, with more complex states resulting from a combination of the former and neutrality denoting a lack of emotion. The second subgroup instead uses a multidimensional approach to define emotion. Usage is largely elicited by a need to analyze the qualitative nature of emotional states, which is often neglected or only partially described by discrete models. To define the full spectrum of emotion, Lang [75] initially proposed the qualities of valence and arousal, respectively ranging from negative to positive and low to high. This model is depicted in Fig. 4.2A, additionally showing the approximate placing of discrete archetypal states within the valence-arousal dimensional space. As more complex states or states with similar characteristics become harder to distinguish (e.g. anger and frustration, both of negative valence with high arousal), other dimensions have also been proposed to model emotion. To exemplify, Mehrabian [76] introduced an additional *dominance* factor to the emotion representation space, ranging from completely dominant to highly submissive to note the feeling or lack thereof control over situations or other people. While the classification of emotion through categorical models may be more straightforward, limitations of oversimplification when dealing with more complex emotional states



(A) 2D Valence-Arousal Space



(B) 3D Valence-Arousal-Dominance Space

Figure 4.2: Juxtaposition of archetypal emotional states from discrete modeling over 2-dimensional (up) and 3-dimensional (down) space.

occasionally warrant translation to multidimensional approaches. A 3D model as shown in Fig. 4.2B, may also allow for better classification of states not clearly separable by 2D valence-arousal standards. Nevertheless, it should be noted any of these types of models are valid and their usage is highly dependent on the desired real-world application.

4.3 Epistemic/Achievement States

By now, the role of emotion in learning through conditioning and reinforcement should be clearer. In addition to this, there is also a set of emotional states that serve as major drivers of the knowledge acquisition process, dubbed epistemic when motivating critical reflection and inquiry, or achievement emotions when being a direct consequence of success or failure [77, 78, 79, 80]. When prompted with a type of (complex) data previously undealt with or when confronted with information contradictory to current knowledge, we enter an epistemic state of confusion. This can instigate engagement and knowledge exploration, so long as the source of that confusion is resolved by this incitement [81]. Surprise may be manifested under high PE or similar conditions to confusion, and has been observed to increase attention on what triggered it [82]. This is possibly due to the reaction underlying the feeling of surprise, which may or may not be enjoyable to the subject. When positive, the subject can seek to reproduce the same feeling by evaluating more data of the same type or even exploring further unrelated information, demonstrating curiosity. A similar but contrasting result occurs with the state of pride. This achievement emotion seemingly stems from a sense of accomplishment in laborious tasks, requiring some level of ability, and has a rewarding effect [83]. Thus, it is a way to sustain learning and can correlate positively with motivation and knowledge exploration. These learning characteristics would be highly beneficial in the context of DL, which by itself is currently lacking in terms of motivation and autonomy. As such, it would be useful to study if these correlations, observed in common human behavior, would transfer to implementations of learning algorithms.

4.4 Contagion and Empathy

In a social setting, the emotion of an individual directly or indirectly impacts that of the surrounding group, and vice-versa, in a process known as emotional contagion [84]. This has been observed not only in humans but some animals as well, as individuals display increasingly confluent behavior during an interaction, and even in digital or online settings [85], where emotional expression is achieved through a different medium altogether. Arousal, a key component of emotion, has been identified as a major catalyst of emotional contagion [86] as it is excitatory to peers and likely a consequence of aroused behavior being notably apparent. When positive, this emotional transfer can translate into improvements in work performance, coordination, and group cooperation, as has been observed in sports from team member to team member and work-related scenarios where the emotional state of workers converges to that of people in leadership roles [87]. Studies that have observed this additionally report the positive effect that subject perception of ongoing positive emotional contagion, either personal or in others, can itself boost the effectiveness of the transfer as feedback. Naturally, the advantages of contagion may become drawbacks should the displayed emotion be negative. This remarks the importance of not only using emotional contagion as a way to boost desired behavioral aspects but also taking care to prevent a negative state from spreading in undesirable settings.

Emotional contagion is not far from the concepts of cognitive and affective empathy. These are respectively defined as either a conscious attempt at understanding a peer's state and considering it for ourselves [88], affected by our perspective, or one's emotional reaction as a consequence of perceiving or predicting the emotion of another individual [89], which causes mirroring or adoption of an affiliated emotional state. Furthermore, empathic interactions can foment benefits similar to those caused by emotional contagion, as influence over a person's state will also affect their interest and motivation levels. In addition, empathy promotes greater social development and bonding with peers, aspects that are fundamental for a healthy lifestyle. Nevertheless, there is a difference related to the innate nature of emotional contagion which, despite also being a feature of empathy, does not characterize it fully. This is because empathy as a skill is generally considered to be learnable and a product of the social environment humans live in [90]. Not surprisingly, empathy has been observed in HRI [91] which is useful for building rapport and improving acceptance of a growing preponderance of robotic and AI assistants in society. Likewise, the efficacy of AI and social robots as agents of emotional contagion has also been observed [86], showing the potential of these concepts to be integrated into AI applications.

Part II

Emotional Competence

Prologue

Emotional competence here refers to a complex repertoire of skills expected of human companions, encompassing identification and perception of states for situationally adaptive and appropriate behavior, such as awareness of emotional communication within relationships and the ability to categorize and discern subjective emotional experience from its active expression in peers [92]. This is because, as explained previously, emotional activations in the brain always impact actions and their elements, either directly or indirectly. Thus, incorporating emotion in responsive behavior, for instance in the form of empathic involvement, is necessary to foment trust and build successful relationships with our peers, given our social nature. Thus, doing so also constitutes the emotional competence of an individual, as the capacity for emotion recognition and emotion expression are naturally correlated.

The increasing prevalence of robots and other user-aiding agents in society [93], amplified by an ambition to avert uncanny feelings during their use [94] points to the integration of emotional competence as a strong possibility for the future of social AI. Not only that but assistive artificial systems also boast greater user acceptance [95] and can account for shortcomings in the care of people with developmental disorders (e.g. ASD) [96], when conveying emotion. Not surprisingly, techniques must be explored first concerning conditions set for the accurate perception of emotion in artificial agents [97]. This refers to the detection of user emotion and adequate responsive expression, which allow AI research to then focus on the adaptive and customizable activity that is becoming a necessity for users.

This part of the thesis addresses two topics of emotional competence respectively related to the importance of considering the correlative nature of emotional states during their analysis, and with the benefit of learning and integrating human social traits (namely empathic involvement) in artificial agent behavior during interactions. While the former may help reduce the effect of missing information during user state analysis and is shown to improve emotion recognition rates, the latter posits a way for robots to learn better social etiquette via user feedback, concluding how the usage of that new knowledge can ease integration and even foster user preference for agents capable of context-appropriate behavior. The impact of these works could also extend to emotion regulation, as perception improves and an agent's ability to adapt towards a certain user emotional goal becomes possible.

State-of-the-art

As mentioned, emotional competence is largely based on two major fields of research. Emotion recognition aims to identify the state or states of a solo or group of users by evaluating patterns displayed in physical/social manifestations of emotional states. Techniques in this field typically involve some form of signal processing methodology combined with DL and deal with one to several modalities to situate data in emotional space based on psychological models of emotion. This can follow either a categorical or dimensional approach, as mentioned in section 4.2. As reviewed in [23], the visual modality and facial emotion recognition (FER) specifically arguably get the most attention, with noteworthy recent works including [98] where expressions were modeled as trajectories on a Riemannian manifold, and [99] where authors augmented texture images with information from local geometric descriptors of their respective 3D meshes. Similar approaches have also been presented for body movement analysis [100]. In terms of speech emotion recognition (SER), novel techniques include mappings to discriminant projection subspaces [101] and acoustic space being partitioned for phoneme posterior probability to highlight emotional relevance [102]. As for other areas, research is comparably reduced as challenges become progressively specific. For example, emotion recognition in physiological signals has attempted to assess the adequacy and explore correlations between EEG channels [103, 104]. In text-based methods there has been focus on semantic analysis with knowledge transfer between different label sets [105]. Regardless, there appears to be a trend in emotion recognition research to focus on refining available data with information potentially relevant for a more holistic perspective or developing novel representations better suited for emotional analysis.

The field of emotion expression is highly interconnected with recognition as these belong to the same action-perception cycle. Its main objective is to generate physical/social reactions in artificial agents, which users can understand and perceive as emotional (i.e. recognize). While preset mappings between eliciting factors and emotional manifestations have been used as non-cognitive approaches in expression research [106, 107, 108], generative DL and RL approaches have become more prevalent here [23], to provide a broader range of multi-modal responses to varied scenarios. To exemplify the former, in [109] authors generated postures and motion patterns through a GAN, while in [110] a dynamic cell structure ANN optimized robot kinematics to appear emotional. Both architectures were made to observe or perform real interactions with users. In [111] facial emotion expression data shaped the latent space of a calligraphy-generating GAN while [112] demonstrated emotional dialog generation via an autoencoder encoding audiovisual information with attention. From these, it is apparent how emotional expression research employs user imitation of some form. As for RL techniques, hard-wired feedback is used to adapt expression according to reinforcement, as in [113] where an agent's facial grimace and vocalizations were rewarded by the laughter caused in the user. Reinforcement may also be indirect, as when obtained from the evaluation metrics of models pre-trained with user emotional data [114]. Nonetheless, explicit feedback is preferred not only to avoid noise when teaching emotional behavior to an agent [115], but also for adaptation to target user preferences during social contexts. [116] demonstrates this by reporting improved HRI when a robot would adapt its speech characteristics using a DQN according to user ratings provided on the fly. Ultimately, RL methodology boasts a framework to consider user feedback, which is not always straightforward or possible in generative techniques and thus may provide some edge in expression research.

Emotion expression may be interpreted in other experiments not targeting user interaction specifically, wherein agents may also be considered somewhat emotionally competent. For instance, Castro-González *et al.* [117] devised a RL appraisal mechanism to learn how dangerous current circumstances may be for a robot's safety. The latter ends up displaying behaviors similar to animals in fear. Similarly in [118], authors described robot navigation as happy or fearful, when based on a common emotional subspace between behavior and environment reinforcement to learn ideal actions. Word-attentive ANNs coupled with meshed-memory transformers expressed perspectives of abstract emotional content in paintings [119], after learning associations between images and user-generated captions. A paradigm was developed that enabled the translation of infant-directed speech features to other modalities, modulating the expression of emotion through these other means besides speech [120]. While these works may be deemed more genuine in terms of true emotional expression and competence, there is controversy largely caused by the topic's complexity [121] and a lack of objective conventions for emotional authenticity.

On another note, standalone models of emotion for recognition or expression may be combined and are occasionally integrated into the drive systems of social robots, such as Sophia [122] and Geminoid [123]. These then form action-perception loops where detected user emotion may, in part or as a whole, constitute the reinforcement signal provided to an agent for behavioral adaptability. In turn, the agent's expressive reaction should elicit another emotion in the user, which the system then detects and repeats the process. Not surprisingly these exact steps, namely state acquisition, analysis, and understanding, followed by adaptive interaction, have also been outlined as a basis scheme for the development of those same emotional HRI and drive systems [124]. Nevertheless, the third step is usually the least robust given how often adaptability pertains to a small set of scenarios while considering more leads to increased complexity and correspondingly decreased performance or interaction quality.

Computational models of emotion represent another section of emotional competence research, not as prominent as others already described. In general, these models are employed as simulators of theories on biological emotional functioning and initially originated from a psychological standing. Contrarily, recent research is more associated with objectivity and neurophysiology, defining characteristics of emotional elicitation and perception in the form of stimuli [23]. The SHArE [125] and HED-ID [126] frameworks exemplify this specifically, focusing on aspects such as valence, arousal, and duration to determine emotion outputs. The lack of popularity here may be related to noted low interaction safety [127] and significant implementation complexity, as well as occasional ethical concerns [128]. These authors have also argued for the development of these models to be more domain-independent and focus on more modulatory associations between appraisal and emotion intensity. This entails requirements analysis and, from a software engineering standpoint, researchers are advised to first identify relationships and data flow between components of an emotion theory when designing computational models for it [129].

All these works have in common the fact that they presented some form of AI with a degree of emotional competence. Yet as mentioned, agreeing on the plausibility of artificial emotional competence is a complicated issue. Thus, requisites for plausibility could be outlined based on human experience of emotional competence, for future research. For instance, agent transition between emotions should be contextually adaptive and smooth/fluid instead of abrupt, as adopted by current approaches. As progress is made, other requisites such as meta-cognition [130] and reaction-action cycling [131] should constitute new objectives. Respectively, these would provide agents awareness of their emotional intelligence plus the ability to control it as well as benefit from its automated facets. As postulated in [23], these traits could enable greater cognitive control for emotional AI agents, by operating proactively and reactively both simultaneously and independently [132], as humans do. Ultimately, user credibility of emotional competence is still a far-out accomplishment and current research must, for the foreseeable future, focus on resolving smaller issues as the ones described previously. This could provide further validity to empirical studies of emotional competence presented by Cognitive Neuroscience, which are based on theory of mind and intention understanding.

Chapter 5

Fuzziness of Emotion

Contents

This chapter presents an experiment that combines DL and fuzzy clustering to support the understanding that emotion analysis would benefit from considering a fuzzy rather than discretized emotional distribution model during recognition. The proposed approach encompassed developing a neural network layer that implements a fuzzy clustering algorithm to fuzzify the emotional features received from upstream layers, so classification is instead based on correlations between states. The aim was towards SER and testing employed four standard emotional speech databases. Obtained results surpassed other SER techniques and were presented in [133], from where this chapter's content is adapted.

5.1 Context

Emotional intelligence is a logical future feature of AI assistive services, given how these are expected to provide support [134] and appropriately empathize with user limitations [135]. This also generalizes to social HRI, where understanding users is key for agent acceptance. To endow machines with this competence, however, it is first required to provide some degree of context awareness through recognition of the emotional states displayed by users through several modalities including speech and prosody, whose variations resonate with emotion. This is because the flow of emotion-related activations over speech-processing neural structures induces changes in intonation and prosody [23], similarly to the introduction of noise over data or the convolution of two separate signals occurring in tandem in the same locale.

While varied models of emotion have been proposed, as previously described, the large majority of works about SER adopt categorical alternatives and DL [23]. To exemplify, recent works include [136], where speech features were mapped to a discriminant projection subspace and classification of states was performed based on a distance metric, or [137], which proposed classifying emotion based on the distributional structure of combined labeled and unlabelled speech data, learned by a generative adversarial network. In [138], authors hard-clustered utterance sequences using a bidirectional LSTM-CNN to recognize states from emotionally dominant features. While these approaches are interesting, simpler techniques have achieved similar results, such as [139] where mel-frequency cepstrum coefficients and modulation spectral features of emotional utterances fed to a small RNN were enough to recognize the respective state. In [140], authors extracted eGeMAPS features and fed these to a Deep Belief Network [141], obtaining accuracies in the high 70% mark despite using few target-corpus instances for training. Regardless of approach, these works fail to consider the correlations between archetypal emotional states that critics have pointed out as missing from categorical models [142], one of which is exemplified in Fig. 5.1A. Transitions between emotions clearly demonstrate this issue, given states are inherently fuzzy and overlapping in real life, more closely resembling the adapted model in Fig. 5.1B, but are limited by sharp divisions in categorical models. This characteristic resonates with fuzzy logic, where boundaries between classes are blurred and evaluated quantitatively or qualitatively, most commonly involving some degree of overlap to better approach the nonlinearity of data. Consequently, data instances may comprise several distinct classes with corresponding degrees of membership or their examination may consider correlations between one another before assigning them labels. A direct benefit of this stems from how these correlations can help fill in information missing from instances of a particular state if common features are shared with other states of lower membership. Approaches have been proposed for FER which employ fuzzy logic to attain these advantages. For example, [143] implements two layers in a neural network for calculating matching degrees of incoming emotional features and uses these to determine the weighting respective of each fuzzy rule employed in the classification averaging, dubbing this fuzzy neural network. A similar technique is used in [144], though with a wavelet network layer in between. In [145], authors adjusted the weights of an emotional neural network using fuzzy c-means clustering combined with a genetic algorithm. Contrastingly, fuzzy logic remains largely unexplored in SER, despite prosody varying similarly for different emotional states. Plus, works reporting improved results over non-fuzzy approaches in the related field of FER add up to the validity of employing these techniques in SER.

Considering the lack of approaches combining fuzzification and SER, we proposed a new technique based on fuzzification of speech features at an intermediate level of a neural network. At a preprocessing stage, spectrograms of emotional utterances are computed. These are passed on as input to a CNN for emotional feature extraction. Obtained results then undergo fuzzy clustering, before classification. While a similar process has been employed in other areas such as object recognition [146], ours differs in that there is heavy dimensionality reduction and specialization of features before fuzzification. This is justified by the fact that these algorithms tend to fail when dealing with data at dimensions greater than a few dozen. Moreover, our approach is a novel step in SER and allowed us to confirm the suitability of fuzzified emotional features for classification. Testing was carried out over four standard SER databases and results were compared with those of other state-of-the-art works in the field.

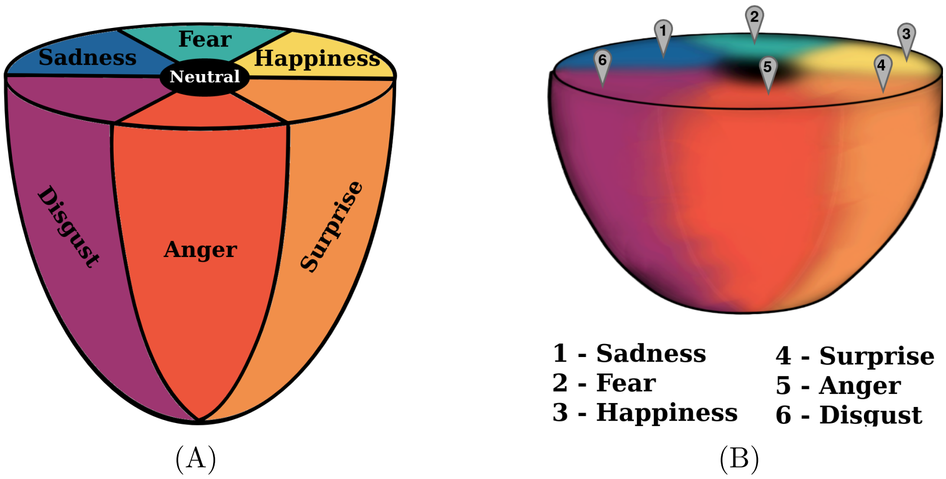


Figure 5.1: Common version of Plutchik's Emotional Wheel [73] (left) and a fuzzy more plausible adaptation of the same model (right).

5.2 Pipeline Overview

As mentioned, the proposed approach encompassed a preprocessing and feature extraction stage, where spectrograms were generated for each emotional utterance, then analyzed by a CNN trained for large-scale speaker recognition and refined by a newly trained multi-layer perceptron. This was followed by the fuzzification stage, which employed the fuzzy c-means algorithm to cluster the extracted speech features, with the number of clusters greatly surpassing the number of classes. Finally, membership degree vectors were classified into archetypal emotions by a perceptron. This pipeline is depicted in Figure 5.2 for easier understanding. Each of these stages is overviewed in the following sections.

5.2.1 Preprocessing & Feature Extraction

To retain as much emotional information as possible, utterances were analyzed in their raw spectral form. This was done by computing spectrograms, representations of energy variation at different frequencies over time which embed prosodic variations (e.g. pitch, tone) as visual characteristics. Specifically, a sliding Hamming window with $25ms$ of width was applied to each considered audio clip, with a step of $10ms$ for smoothing, after which the corresponding discrete Fourier transform was obtained. The respective means and variances were also normalized at every frequency bin (i.e. the intervals between samples in the frequency domain) of the spectrum so all utterances would be in the same range of values before being analyzed. No additional preprocessing was performed over the data.

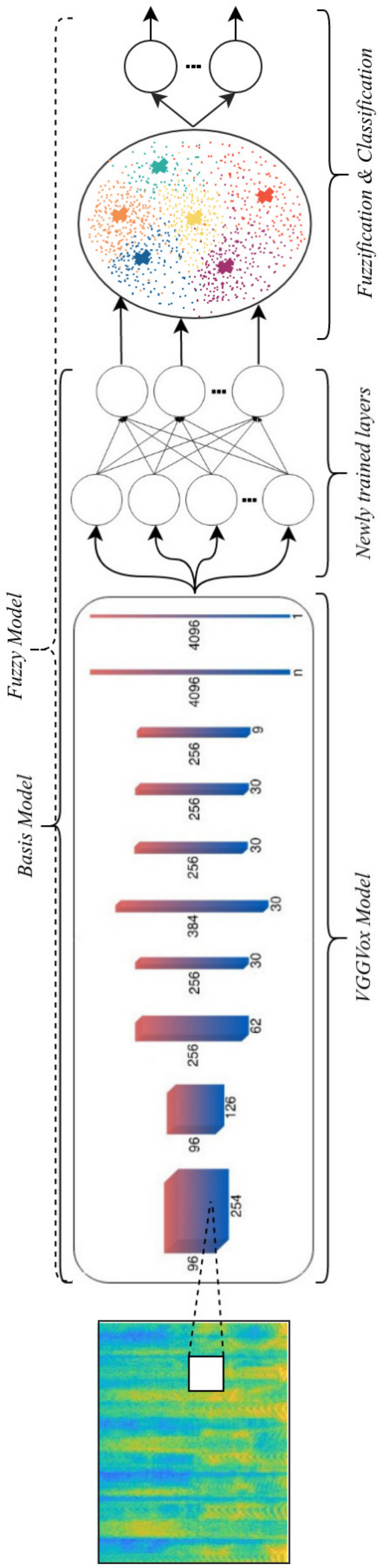


Figure 5.2: Structure overview of the proposed model incorporating a fuzzy layer.

Given the bi-dimensional nature of the preprocessed spectral data, CNNs were an evident choice for its analysis and feature extraction. However, the unavailability of big data in emotional speech precludes acceptable performance by directly trained DL architectures, meaning the obtained features would not be very useful. Nevertheless, other speech-related tasks exist where data is abundant with models trained on them being capable of successfully extracting useful prosody information. Thus the VGGVox model [147], a large-scale CNN specifically designed for the analysis of audio in its spectral form and respective recognition of speaker, was pruned and its upstream layers employed for feature extraction. Being a VGG-M architecture, the model is composed of a sequence of convolutional layers each of which is followed by a pooling operation. As for the section of VGGVox employed in our approach, 6 levels of the convolution-pooling sequence were kept while the rest of the network was removed. This is because at this level the network is capable of extracting features useful for speech-related tasks, while not yet being fully specialized only for the task of speaker recognition [148]. After this pruned VGGVox model, two fully connected layers were directly trained to further specialize the features extracted for emotion recognition, using the emotional speech data available. Moreover, the variable input length capability of VGGVox was also implemented in the final layer of this section, as it is desirable for audio analysis since the model becomes invariant to temporal position but not frequency [149].

5.2.2 Fuzzification

For the fuzzification part of the proposed architecture, the developed layer employed the fuzzy c-means method [150] to cluster embeddings according to their similarities. This algorithm was selected given that its simplistic approach makes it ideal for establishing a baseline of fuzzy logic techniques in SER, which is nonexistent, and enables subsequent comparison with more complex techniques such as possibilistic-c-means. Considering an instance x_j made up of L features $\mathbf{F} = (f_1, ..., f_L)$ and given a pre-set number of clusters K, the algorithm starts with $C = \{c_1, ..., c_K\}$ cluster centroids, initialized according to a uniform distribution, which are then iteratively updated to better match the data. These are computed through a weighted average of all points:

$$c_k = \frac{\sum_{j=1}^{N} \omega_{jk}^{m} \cdot x_j}{\sum_{j=1}^{N} \omega_{jk}^{m}} \tag{5.1}$$

Here, $m > 1$ is the fuzziness intensity of a cluster which asserts how much it may overlap with others in proximity. This was kept at a default value of 2. The membership degree of embedding x_j in cluster k is represented by ω_{jk} and obtained through:

$$\omega_{jk} = \frac{1}{\sum_{i=1}^{K} \left(\frac{||x_j - c_k||}{||x_j - c_i||}\right)^{\frac{2}{m-1}}} \tag{5.2}$$

This implementation was made a type-1 fuzzy system since, while a level of uncertainty could have been associated with ω_{jk}, that was not added due to the potential introduction of noise and excessive homogeneity in classification caused by too many cluster centroids being considered. Minimization of the weighted sum of all possible instance-centroid pair squared norms was performed until a maximum number of iterations was reached or centroid updates became negligible, to find optimized locations. The maximum number of iterations was set as 150, while the centroid threshold was set as 0.00001. Once trained, this new fuzzy layer received the emotion-specialized feature vectors from the CNN+perceptron basis model and provided a set of membership degrees ω respective to each cluster and its centroid. Finally, these membership vectors were progressed through a fully connected layer, likewise newly trained with the available emotional speech data, for final classification.

5.3 Experimental Results

The experimental section designed for this system first encompassed determining an appropriate number of clusters to be used for fuzzification. This is because there is an ideal range for this parameter to benefit model performance. Secondly, an ablation study was carried out to assess the utility of the fuzzification layer for improving classification. The databases considered were EMODB [151], EMOVO [152], SAVEE [153] and ELRA-S0329 [154] given their widespread use in SER, making comparison of results easier. Each database uses a different language and considers 7 emotional states with the common ones being anger, happiness, sadness, fear, disgust, and the neutral state. Boredom is considered by EMODB, while the rest adopt surprise instead.

It was hypothesized that considering a group of emotions intermediate to the main set of archetypal states could help analyze ambiguous utterances. Naturally, by making this set of intermediate states larger, correlations between states become increasingly more prevalent until a certain point where individual state information becomes overly diluted. For this reason, increasing quantities of cluster centroids were tested until a performance drop was observed. This was achieved by training and testing the fuzzy model with the SER data from all databases using 10-fold cross-validation and 30 epochs. As can be observed from Table 5.1, there was a steady increase in performance from 10 until 200 clusters, at which the expected drop occurred. For this reason, the fuzzy model employed in the following part of the study was implemented using 150 clusters.

Table 5.1: Fuzzy layer performance variation by number of clusters, using 30 epochs and 10-fold cross-validation. Accuracy values are percentage points.

Clusters	10	25	50	100	150	200
Acc (Std)	63.75 (4.00)	64.12 (5.32)	65.42 (8.25)	67.10 (4.82)	68.93 (5.35)	63.33 (6.68)

For the ablation part of this study, we intended to demonstrate how the fuzzy layer can account for shortcomings in emotional data and improve the performance of an SER model by taking advantage of inter-emotional state relationships. The main goal was to obtain a performance increase similar to reports from FER works, where fuzzification has been successfully incorporated. To this end, the same training process and hyperparameters were applied to the fuzzy model described thus far (see Fig. 5.2), and to a basis model with the same added fully connected and classification layers albeit abstained from any data fuzzification. This process considered each emotional database individually and followed a 10-fold cross-validation policy for training and testing, systematically increasing the number of epochs. The obtained results are detailed in Table 5.2 in terms of mean accuracy and standard deviation, with respect to the policy used. Results from three other state-of-the-art works are also reported for comparison. As can be observed, the fuzzy model performed better than its non-fuzzy counterpart in all but the 30-epoch test with the EMOVO database. Additionally, the best results obtained with the fuzzy model were on par with or surpassed those of other works in the field, again except for the EMOVO database.

Table 5.2: Comparison of accuracy results between the non-fuzzy and fuzzy models trained over 30, 50, or 100 epochs and tested based on 10-fold cross-validation, and against other state-of-the-art techniques. Accuracy values are percentage points. The mean of performance increases across databases is shown in the rightmost column.

Epochs	**Model**	**EMODB**	**SAVEE**	**EMOVO**	**S0329**	**Mean Gain**
		Accuracy μ (Standard Deviation σ)				
30	*Non-Fuzzy*	67.05 (5.61)	58.54 (8.08)	51.02 (8.02)	88.38 (3.47)	1.66
	Fuzzy	68.93 (5.35)	61.25 (5.88)	50.85 (7.19)	90.61 (3.03)	
50	*Non-Fuzzy*	71.94 (5.30)	63.54 (4.61)	55.42 (4.55)	88.97 (2.97)	1.65
	Fuzzy	74.17 (5.71)	66.04 (6.27)	55.93 (3.26)	90.31 (2.11)	
100	*Non-Fuzzy*	76.62 (8.41)	68.54 (2.98)	62.20 (7.42)	90.91 (3.19)	1.71
	Fuzzy	78.48 (7.86)	71.04 (4.76)	64.07 (7.72)	91.51 (2.68)	
Other Works	*Kerkeni* [139]	69.6	-	-	90.1	
	Latif [140]	72.4	56.8	76.2	-	
	Sidorov [155]	74.6	63.8	-	-	

5.4 Discussion

Results in Table 5.2 clearly highlight an increased performance when employing fuzzification in the DL model, regardless of the database when completing 50 or more epochs. Based on this, we can conclude how archetypal emotional states do share common traits and are frequently correlated with one another when expressed vocally. As for the lack of improvement over EMOVO with 30 epochs, this may imply a high level of emotion heterogeneity in the used language (Italian) which hinders the classification of fuzzified data since emotional states are less or not at all correlated. Regardless, improvement becomes apparent for EMOVO as well when employing a greater number of epochs. In fact, this is also true for the tests performed with each database, indicating a trend for performance growth which could be explored in further works. Nevertheless, given how the mean performance gain remains nearly constant despite the epoch number, it can be concluded that this hyperparameter variation does not intervene in the increased accuracy from the non-fuzzy to the fuzzy model. Consequently, these results support the initial hypothesis that fuzzification can be used to take advantage of inter-state correlations to account for potentially missing information and produce features better suited for classification or at the very least increase robustness, similarly to what has been attained in FER.

While the lower EMOVO accuracy in comparison with other works may be similarly due to the heterogeneity of Italian emotional speech which renders fuzzification counterproductive, the increased performance on other databases is expected and follows our initial hypothesis. This additionally motivates studying the effect of fuzzification on other DL architectures as well as more robust fuzzy logic techniques for SER in general, since an improvement baseline has now been established. Finally, given this first approach to fuzzy SER was successful, joint fuzzification of emotional audio and visual features should be explored towards multi-modal emotion recognition and to assess the degree to which two states may correlate and complement one another even when expressed in distinct modalities.

Chapter 6

Empathy in Social HRI

Contents

This chapter describes an experiment meant to assess the impact of affective empathy on human-robot interactions. The process involved training a DL module to perform facial expressions appropriate to each emotional state, via reinforcement learning. A real interaction scenario was then designed, with a set of participants being asked to engage with a robot endowed with said trained module. The experiment was designed and carried out in collaboration with colleague Alessandra Sorrentino, PhD., from the BioRobotics Institute of Scuola Superiore Sant'Anna in Pisa, Italy, and presented in [156] and [157], from where this content is adapted.

6.1 Context

Understanding how humans perceive robots and the impact of certain behavioral traits in the quality of HRI is a natural step towards bettering engagement and user acceptance. One way to contribute to that objective involves integrating the affective characteristics of empathy and emotional contagion in robot frameworks. To mimic the learnable nature of the former in social robots, computational empathy emerged as a field wherein techniques are developed for agents to be capable of behaving empathically with their users and benefit from it. Naturally, techniques also involve contagion, considering its innateness, which indirectly benefits interactions.

HRI studies have posited the importance of artificial agents understanding the human psyche so a rapport can be built with their users [158], with some researchers even arguing that empathy could be achieved in AI if agents were endowed with an artificial nervous system and pain induction [159]. Plus, a clear preference for empathic agents has been observed in cooperative tasks [160]. Furthermore, empathic behavior has been shown to boost user trust and fondness in human-robot relationships [161]. This is likely a consequence of the comfort increase and stress decrease advantages of empathic interactions which transfer to HRI [162]. Empathy can be perceived through as many modalities as emotion can be expressed or recognized in, which by themselves generally depend on both physical and behavioral attributes. While many studies fixate on improving human likeness in robots to showcase empathy, this study focused on a behavioral aspect of the trait as well due to this being a major component of empathy in real life [163]. Hardware restrictions and a more challenging perception of user state [164] are the probable causes of the reduced number of works being developed for empathic behavior in real robots. Thus, even though related literature here could encompass works where virtual scenarios serve as proof of concept, this study focused more on implementations that have been deployed and tested in the field. For instance, in [135], authors analyzed whether elderly users perceived empathy from a NAO robot performing affective acts based on utterances, body movement, and color variation in line with the state perceived from user speech, facial expressions, and gestures. Recognizing and imitating user facial expressions was found to benefit a robotic head in terms of how users perceived it during dialogue in [165]. Similar effects were observed in [166], as users rated interactions more positively when the mimicking of emotional gestures by a robot head was more pronounced. The work in [167] describes a mobile virtual agent which varies its facial expression, movement and sound effects to mimic the perceived emotion of autistic children. This models interaction patterns for distinct personalities and thus foments engagement. Bagheri *et al.* reported increased comfort and confidence in users interacting with a Pepper robot which matched detected user emotions and personality in its utterances, learned via a contextual bandit RL approach. Despite not mentioning empathy explicitly, [168] teaches a robot to express the emotion of its user by using human feedback to minimize the error between the target state, perceived visually, and a perceptron-chosen robot expression. In [169], the same authors further their work by using an actor-critic RL technique to associate robot expressions to emotions described by a growing-when-required network in a story-telling scenario, leading to a more natural interaction.

The main issues identified in the reviewed literature are related to robot behavior, for instance, via vocalizations or facial expressions, being hard-wired to rather than learned from user emotion. Moreover, a part of the studies also lacks an adequate interaction scenario, reporting results based on a bird's-eye evaluation of performance without much user-robot interaction. In our experiment, eye and mouth configurations are associated by feedback to archetypal emotional states expressed by human users to mirror them and simulate affective empathy, in an initial stage. A DQN was developed as the learning approach to teach the robot an adequate action-selection policy when it came to expressing emotions via its facial configuration. This was elicited by the high-dimensional state space associated with mapping facial expressions to emotional states. Our methodology is most closely related with [168, 169], given RL was used to determine expression. Nevertheless, besides employing the same set of archetypal emotions, we developed a distinct learning approach and considered a fairly larger set of facial expressions that stem from a combination of eye and mouth configurations learned by distinct branches of a model. Thus, the algorithm is meant to learn an adequate combination of physical attributes rather than merely selecting the most appropriate expression from a list. Moreover, this can equate to the same eye/mouth configuration being mapped to different states or vice-versa, which is more plausible in a real scenario. Finally, in the second stage, the robot's learned empathy module is deployed in a short dialogue scenario with real users to determine whether empathy was conveyed.

6.2 Learning Stage - Mirroring Optimization

As mentioned, this stage focused on training a model to map emotional states with combinations of eye and mouth configurations. Emotion was viewed as categorical, being either neutrality, happiness, sadness, anger, surprise, fear, and disgust, as this is the most commonly employed approach and archetypal states are typically innate and readily apparent [23]. Regarding the artificial agent, the GrowMU robot [170] was used, which allowed us to define the eye and mouth configurations based on its available modules as well as on the Facial Action Coding System (FACS) [171], which isolate specific points of the human face involved in affective expression. This resulted in 13 mouth and 4 eye configurations, which combined makeup 52 possible facial expressions, shown in Fig. 6.1.

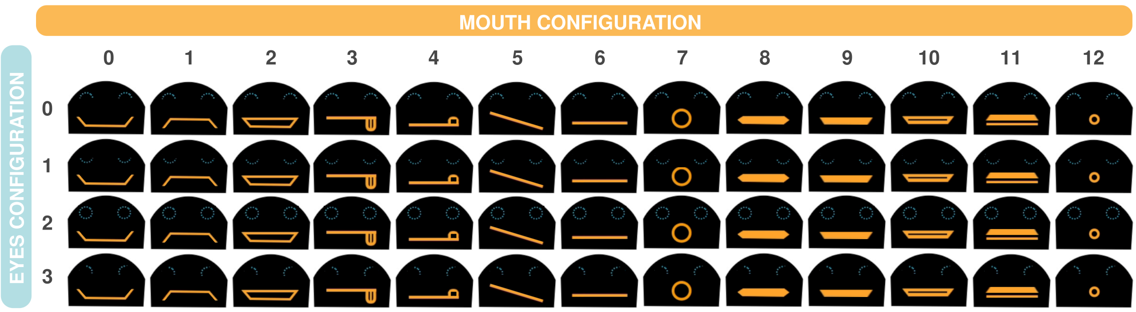


Figure 6.1: Possible combinations of eye and mouth configurations to form the robot's facial expression. As presented in [156].

6.2.1 RL Outline

Following the formulation of ϵ-greedy DQNs in section 2.3.1, the environment's action-space A envisioned for this experiment encompassed the 52 combinations of eye and mouth styles. Also, the state-space S was made up of two disparate states representing empathy or lack thereof, respectively whether the robot-generated facial expression matched the detected user emotional state or not. Based on this framework, a reward system r_t was designed to motivate the agent at each step t towards demonstrating what its implementation considered to be empathy. Specifically, the system rewarded the agent one of two possible constant values, $r_t = 2$

whenever it moved from the non-empathic to the empathic state and $r_t = 10$ should the agent decide to remain in the empathic state. This is because it should be the robot's objective to demonstrate empathy as much as possible towards its users during an interaction. Consequently, moving to or remaining in a non-empathic state resulted in a null reward $r_t = 0$ being provided to the agent. This reward was not made negative as an agent being non-empathic was not necessarily considered to be incorrect behavior, thus not requiring punishment.

In terms of the neural network architecture designed as the DQN of this experiment, its root section was composed of two fully connected layers and received as input the concatenation of both the current state s_t and user emotion. The architecture was kept simple to avoid convergence issues stemming from larger function approximators. From this section, two further fully connected layers were specialized, with one corresponding to eye and another to mouth features. The processing of these features was implemented in separate layers in conformity with biological neural processing, where upstream information stemming from limbic structures is passed on to outward cortices specialized for each facial function. Finally, each branch outputs a configuration out of the 4 or 13 available for either the eyes or mouth, respectively. The RL environment and neural architecture designed for the robotic agent are both depicted in Fig. 6.2, for easier understanding.

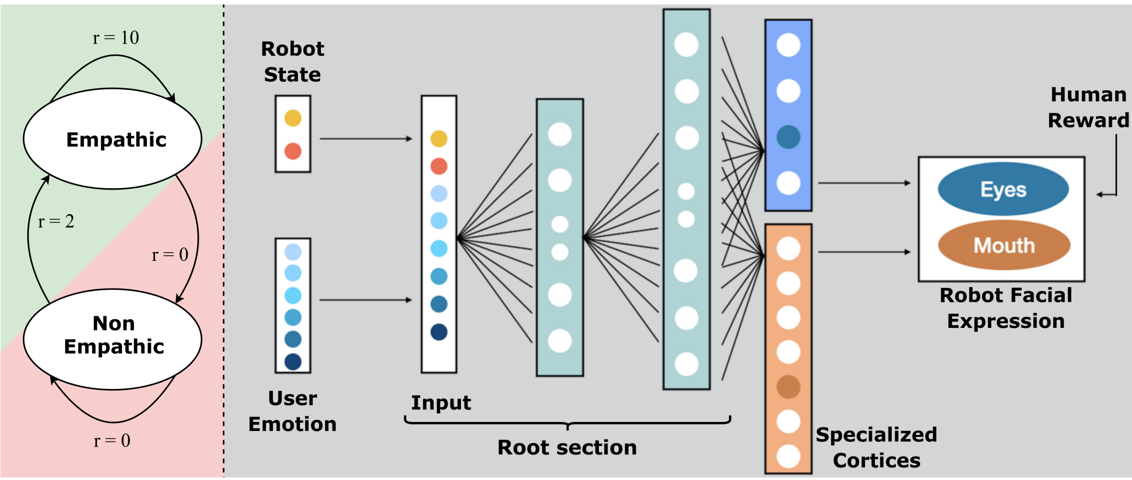


Figure 6.2: Environment (left) and neural architecture (right) designed for the RL robotic agent to demonstrate empathy by matching user emotion with its facial expression. Adapted from [156].

6.2.2 Training Process

An agent requires quite a large amount of samples of the environment if it is to learn an adequate policy using DQNs. While it is possible to obtain these samples via physical interactions with one user at a time, implementation over a single robot would mean these would consume an amount of time infeasible with the completion of this experiment. For this reason, an online platform was developed and connected to the RL algorithm via a WebSocket, so the artificial agent could learn from several users simultaneously. An exemplary screenshot of the platform is shown in

Fig. 6.3. Based on a randomly generated target emotion and the agent's current state concerning a user, eye and mouth configurations were obtained by the DQN and replicated in the user's browser identically to how they would be shown by the physical robot. To determine the reward value to award the agent, the user was then asked to classify the demonstrated facial expression as either coherent or incoherent with the target emotional state, according to their own opinion by clicking a button. This rating was stored locally in a log file and used by the algorithm to perform an iteration according to standard DQN procedure.

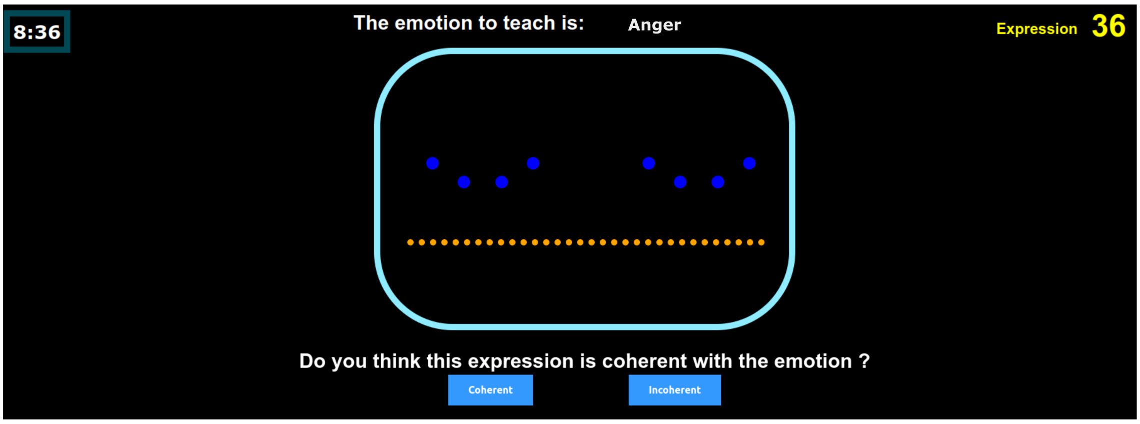


Figure 6.3: Screenshot of a training session using the developed online platform. The target emotion is shown on top, based on which the user should classify the simulated robot facial expression as coherent or incoherent. The bottom left corner shows the elapsed session time while the right corner shows the number of expressions evaluated so far.

In order to gather a large number of participants in this stage of the experiment, the URL of the platform was disseminated by various means such as e-mail blasts and social media. By accepting the invitation, participants were shown a brief tutorial detailing what to do in this experiment, as well as a short overview of its objective. Subsequently, the user would have to fill out a short socio-demographic form, for statistical purposes. Following this, each participant was directed to the training page, where facial expressions were evaluated, after consenting to their anonymous data being stored and used for research purposes. Once there, a training session would last as long as the user wanted or until a maximum of 10 minutes, at which time the user could restart the process and perform more sessions or end their participation. Moreover, the target emotion would change every 250 user responses, corresponding to the number of steps per learning episode, with a pop-up alert informing the user of that change. With this process the training of the RL model was continuous and copies of it were stored locally at periodical intervals so that the model could be accessible to the physical robot at different stages of training.

6.2.3 Results

Based on the short socio-demographic survey, a total of 105 participants were registered, with 46 females and 59 males at an average age of 32.01 years ($\sigma = 10.06$). In terms of educational background, a small percentage of participants had high-school qualifications or lower (5.83%), with the remainder reporting either a bachelor's degree (13.59%), a master's degree (60.19%) or higher (20.39%). Though most of the participants were either Italian (58.25%) or Portuguese (26.21%), other nationalities were also registered from Asia (5.83%), Europe (6.8%), South America (1.94%) and North Africa (0.97%). All in all, participants provided an average of 155.6 responses each ($\sigma = 92.95$), with the total amount of facial configurations generated by the model and evaluated by users being 22251. The total amount of expressions generated for each emotional state is shown on the bottom row of Table 6.1. Additionally, the response log showed that each facial configuration was explored several times for each emotional state, meaning the RL algorithm adequately explored its environment while training. *Coherent* ratings of facial configuration to emotional state composed 42.51% of all responses provided by participants. In fact, referring to the matrix of expressions shown in Fig. 6.1 in terms of $[eye, mouth]$, the combinations $[1, 1], [1, 6], [1, 11], [2, 1], [2, 2], [2, 6], [2, 11], [2, 12], [3, 1], [3, 6], [3, 11]$ received the most *coherent* ratings from the participants, based on a threshold of 200 *coherent* responses. The ratings of these 11 configurations, normalized by the total number of expressions per emotion, are also shown in Table 6.1 along with the ratio of *coherent* to *incoherent* feedback per configuration. As can be observed, each state exhibits 1 to 3 configurations with higher ratings than the remaining 8 to 10. In addition, 9 out of the 11 configurations boast a ratio higher than 1. Consequently, it is plausible to assume configurations with higher ratios would be more associated with the emotional states where the respective ratings are also higher if the RL model converged properly.

Considering how episode length varied depending on the number of user responses per session, the cumulative reward was considered unsuitable to indicate model convergence. Instead, since the implemented DQN was saved periodically during training, three instances were selected for analysis corresponding to initial, final, and mid-training copies of the model. These were tested by progressing each emotional state through them and comparing the obtained expressions, reported on the bottom rows of Table 6.1, with the information extracted from the user response log files. While the first model yields only 3 distinct facial configurations for the 7 different states, these become more diversified in the middle and final models indicating convergence. Moreover, some configurations were retained either partially (e.g. mouth configuration for the neutral state, eye configuration for sadness) or fully (e.g. anger state, surprise state) when comparing the middle and final models. The associations made by the latter also closely match the results expected from the previous analysis of the user response log files. Specifically, the expressions yielded by the final model for the neutral, happiness, and anger states correspond to the ones that got the highest *coherent* ratings, whilst the expressions obtained for the sadness and surprise states are fairly close to the highest rating values. The expression obtained for fear somewhat deviates from the ratings metric, though it corresponds to one of the higher ratios of *coherent* to *incoherent* responses. Finally, the expression for disgust was the only one not matching any of the 11 expected expressions, likely as a drawback of using a large function approximator such as a DQN. Regardless, based on these results, it can be concluded the model converged

Table 6.1: *Coeherent* ratings of the 11 facial expressions most rated as *coeherent*, normalized by the total number of expressions per emotional state (at the bottom) and with a ratio of *coeherent* to *incoeherent* feedback shown on the rightmost column. Cells in bold correspond to the associations of facial configuration to emotional state made by the final RL model. Expressions associated with each emotion by the RL model at different training stages compose the bottom three rows.

Facial Expression	**Emotional State**							
	Neutral	*Happiness*	*Sadness*	*Anger*	*Surprise*	*Fear*	*Disgust*	**Ratio**
[1,1]	0.0	0.0	16.4	0.1	0.0	0.0	0.2	2.68
[1,6]	0.7	0.6	**14.0**	0.0	0.0	0.2	0.3	3.06
[1,11]	0.2	0.0	16.4	0.3	0.0	0.1	0.2	1.61
[2,1]	0.0	0.0	0.8	0.1	0.0	5.7	1.0	0.7
[2,2]	0.0	**16.7**	0.0	0.0	3.3	0.1	0.0	0.75
[2,6]	**28.1**	0.5	0.1	0.1	3.8	5.9	1.7	1.13
[2,11]	4.6	0.1	0.3	0.0	14.9	12.8	5.8	1.48
[2,12]	0.0	0.0	0.0	0.0	**13.4**	**3.8**	0.0	2.43
[3,1]	0.0	0.0	0.2	**21.9**	0.0	0.1	12.1	2.9
[3,6]	1.8	0.0	0.1	10.3	0.0	0.2	1.6	1.85
[3,11]	4.1	0.1	0.1	11.8	0.0	0.1	4.5	1.41
Total Expressions	2158	2437	3417	3923	3562	3477	3277	
Initial Model	[1,12]	[2,9]	[2,9]	[1,12]	[1,12]	[1,12]	[2,7]	
Middle Model	[3,6]	[0,6]	[1,1]	[3,1]	[2,12]	[0,12]	[3,1]	
Final Model	[2,6]	[2,2]	[1,6]	[3,1]	[2,12]	[2,12]	[3,12]	

appropriately enough to be employed in the next phase of the experiment.

6.3 Deployment Stage - Interaction

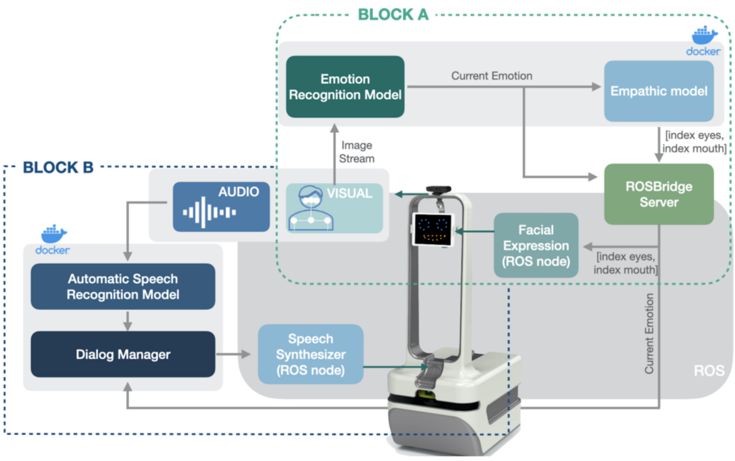
For the sake of completion, this section presents an abridged overview of the experiment's second stage, jointly designed as a testbench of the RL model but mainly implemented by Dr. Alessandra Sorrentino. During an interaction the aim was for the robotic agent, endowed with the model, to match its facial expression with the detected user state, to evaluate the impact of empathy over user perception. Despite the initial plan envisioning a GrowMU robot, a CloudIA robot [172] available at the time of implementation was used.

6.3.1 Design and Setting

The module implemented received input from the visual and auditory modalities, with the former employed for FER of the user via a pre-trained CNN[1] which outputs the 7 emotional states considered in the experiment's first stage. While speech emotion recognition as described in Chapter 5 could have been an alternative here, FER techniques are generally more robust and consistently provide higher accuracy [23]. The user emotion was subsequently employed by the DQN trained previously,

[1]https://github.com/SanjayMarreddi/Emotion-Investigator

along with the current agent state, to compute a facial configuration demonstrative of empathy. This result was passed onto CloudIA's tablet, which displays the chosen configuration to the user during the interaction. As for the auditory modality, the Vosk[2] automatic speech recognition (ASR) model was used to capture user utterances, to which the robot responds as a conversation scenario. This exchange of utterances is managed by a finite state machine, based on what the user says, to determine a response for the robot to vocalize. These vocalizations employ the ROS wrapper of svox-pico Text-to-Speech engine [3], with which the robot's prosody is adapted in terms of tone, pitch, and speed. The communication between the ROS processes controlling the described modules was done using the ROSBridge Server. A diagram of this implementation is shown in Fig. 6.4A.



(A) System Diagram

(B) Non-Empathic

(C) Empathic

Figure 6.4: Overview of the design implemented in CloudIA and experimental setting showing empathic and non-empathic behavior examples.

The conversation between the robot and the user was divided into three levels. The first included welcoming remarks from the robot and basic questions (e.g. "How are you today?") as a way to gently initiate an interaction with a single user. In the second step, the robot actively tried to elicit emotional states in the user by showing videos composed of combined data from the International Affective Picture System (IAPS) [173] and the International Affective Digitized Sounds (IADS) [174] databases. Images corresponding to a particular emotional state were matched with appropriate sounds of the same state and played. Thirdly, an Akinator-styled game was played using the process described in [165], with CloudIA asking a fixed set of questions to guess a character pictured by the user. The robot's behavior during this interaction was set as either empathic or non-empathic (Figs. 6.4C, 6.4B). Users were asked to participate in two interactive sessions with it, one per behavior type with a five-day interval in between. In the first case, the DQN-based facial configuration was active and functioning as described. Moreover, when playing eliciting videos, the robot's utterance sought to confirm and encourage/discourage the emotional state detected in the user via FER, if it matched the video's target

[2]https://alphacephei.com/vosk/
[3]https://github.com/ScazLab/svox tts

emotion. Otherwise, the robot vocalized the video's target emotion as its own. When being non-empathic, the robot's facial configuration was kept neutral and comments were randomized, being either vague (e.g. "wow"), positive, or negative.

6.3.2 Results

All 9 users signed a consent form and were asked to fill out a socio-demographic questionnaire to log user characteristics and experience with robots, followed by the BFI-10 survey [175] to assess the impact of user personality on robot perception, and the PANAS survey [176] to determine their mood at the beginning of the session. After each session, users evaluated their perception of the robot by filling out a modified Godspeed questionnaire [177]. Post both sessions, a short interview was conducted to determine how participants considered robot behavior to differ per session and if so, which of the two they preferred.

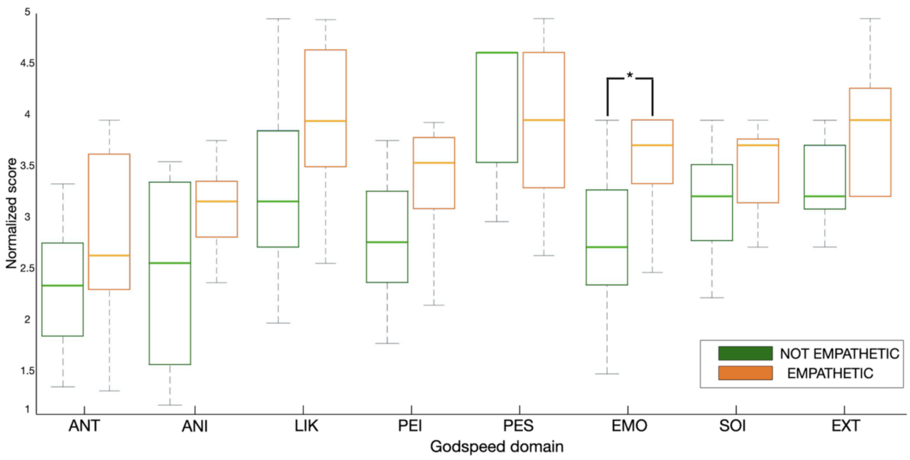


Figure 6.5: Box plot of ratings associated with each Godspeed domain in each behavior mode.

Participant mood was consistently reported as having a discrete presence of positive and a low presence of negative affects at the start of each session. The influence of socio-demographic traits on experimental results was found to be negligible. Results from BFI revealed less conscientious users perceive the robot as safer, more extroverted, and socially intelligent when in empathic mode, while people with lower self-confidence also reported a higher perception of social intelligence in the agent. Despite non-empathic behavior not correlating with BFI reports, PANAS revealed an association between perceived robot intelligence and positive affects in this mode. The Godspeed questionnaire provided the most insight, displayed in Fig. 6.5. Perceived safety (PES) aside, all other domains show a higher score for the empathic mode, indicating users may feel safer with an emotionless robot. Contrastingly, the difference between modes was most significant in the emotion (EMO) domain, meaning the robot was successfully perceived by users as empathic and more humane when in empathic mode rather than non-empathic. This was corroborated by the final interview results, in which 88.89% of users noted CloudIA's facial expression in empathic mode to vary considerably, with 66.67% stating utterances were in line with video content. Altogether, users preferred the robot's empathic behavior described as more expressive and aware over its non-empathic version.

6.4 Overall Discussion

This experiment sought to create an empathic behavior framework to understand how this phenomenon can impact and improve HRI. The backbone of this framework relies on deep learning to detect the user's emotional state and to select an adequate facial configuration for the robot indicative of empathy. While the former was achieved by a pre-trained CNN, the latter employed a new RL model trained via a web-based replication of robot facial behavior. This strategy allowed us to reach a wider population of users and obtain an amount of feedback comparable to several weeks of one-on-one HRI, which would be infeasible. The resulting RL model successfully converged to a set of configurations congruent with the feedback provided by human users, out of 52 combinations of distinct eyes and mouth dispositions. Even though this stage lacked physical embodiment, it could be concluded that robot emotional expression is possible and interpretable by humans. More importantly, given how this computational process resulted in an empathic framework, it is plausible to conclude humans develop empathy, a learnable skill, via a similar but more complex system. Considering how humans are a social species and interpret perceived emotions as positive or negative during interactions, this feedback can serve as a guide for children to adjust their behavior to become more empathic. Consequently, the goal of greater social acceptance is achieved and interactions become easier.

The developed model for empathic facial configurations was integrated into a real robot and employed during a conversation scenario, to assess the effect empathy had over user perception of that robot as well as how user traits would influence this process. Surveying participant opinions via standardized questionnaires revealed a preference for the robot displaying emphatic behavior during the interaction, as this was associated with greater compassion and awareness. This observation shows the importance of enabling artificial agents to adapt their facial expressions to the context they are integrated with, at each moment. Nevertheless, the lower perceived safety by users during empathic sessions may indicate an adaptation period, during which humans distrust agent behavior but which is eventually surpassed. Given how user personality has a discrete effect on robot perception, profiling by an empathic framework may also help reduce this user adaptation period. Regardless, on a broader note, these findings emphasize the significance of not only empathy but general emotional competence in artificial social companions and HRI, as agents become increasingly acceptable the more their social behavior corresponds to that of their users, enabling bonding and improved life quality.

Part III

Adaptability

Prologue

Human and most other animal brains boast numerous organizational and adaptive characteristics that may determine the extent of their behavioral intelligence [178] and consequently provide an evolutionary edge over organisms lacking similar traits. These, as summarily overviewed in Chapters 3 and 4, influence how learning occurs, its efficiency and efficacy, as well as several other variables. Hence, adaptability here concerns the cognitive neural processes related to learning and behavior which enable some degree of autonomy and/or provide an advantage to their host being, when executing some given task. Examples include the neuromodulatory backbone of neuroplasticity and the correlation between certain emotions and exploration which, respectively, can alter the physical structure of neural circuitry to better suit a task and modulate motivation to perform that same task without the need for external intervention.

The term narrow is used to describe current AI methodologies, given their inability to cope with scenarios outside their scope [179]. Recurring issues include lack of autonomy, intrinsic motivation, and plastic structuring. This is differentiated from general AI or AGI, an ambitious ideology of general-purpose systems more similar to humans and versatile towards multi-tasking, which could solve those issues currently affecting social and industrial robotics. Yet, AGI is quite far from realization [180]. A more plausible solution entails emulating and testing the viability of neurophysiological processes in AI methodology so benefits may be obtained for learning and behavior similar to the effects observed in real life [181]. For instance, artificial agents self-motivated by internal processes could decide to analyze data without user instructions, acquiring knowledge during idle time. Moreover, the testing of theorized neural processes in DL architectures could also provide further insight into their veracity and help other fields of research.

This part of the thesis presents two works carried out in an attempt to improve AI adaptability by simulating biological neural processes. First, an emulation of dopaminergic neurons was attempted in ANNs to potentiate connection weights according to the respective changes during training. The goal was to expedite model convergence while also observing dopaminergic behavior when neural networks deal with different data categories (e.g. novel). The second work focused on integrating into artificial agents the modulatory effects that epistemic and achievement emotions have on exploratory behavior. The process involved mimicking underlying neural circuitry, intending to demonstrate how intrinsic motivation can be achieved in AI via emotion integration. Additionally, this work was intended as a corroboration of findings reported in Psychology research, besides providing a path for future experiments correlating emotion and exploration.

State-of-the-art

Emotion is particularly important for adaptability, as is well-documented. Moreover, it is already a significant part of AI research [23], given its role as a reinforcer of efficient learning and adaptive behavior. In RL works specifically, emotion is largely employed in a bottom-up approach (e.g. [182, 183, 184]), wherein emotional traits are implemented in agents as intrinsic boosting of environmental/extrinsic rewards or regulators of state-space. The goal is then to enable new capabilities in agents stemming from appraisal-modulated rather than purely stochastic decision-making [185]. Contrarily, while top-down RL approaches where emotion is emergent or an epiphenomenon are also common [118, 117], its subsequent consideration for agent behavior is considered as emotional competence and thus was overviewed in the previous part of this thesis.

As initially presented by Moerland [186], current emotional RL can be categorized compositely in terms of elicitation, type (either categorical, dimensional, or neither), and function. Thus, some techniques implement emotion as a derivation of agent homeostasis, in turn affected by extrinsic elements tied with state action. This has been demonstrated in [187] and [188], where the variation of robot consumed energy constituted its emotion. Diversely, in [189], [182], and [190], agent emotion stemmed from valence, novelty or uncertainty. This constitutes another form of elicitation termed stimuli appraisal, with agents being motivated by the worth perceived from data features. It should not be confused with approaches where emotion is elicited by the value and/or reward functions of an agent, such as [191] or [192] where emotion was derived from the temporal difference error calculation or difference between short and long-term mean reward entropy, respectively. Finally, elicitation may also be hard-wired, wherein emotion is obtained from sensory input instead of internal agent parameters. In [193], this was demonstrated from decay dynamics, while in [194] user feedback was employed.

Aside from epiphenomenons, emotion can serve several functions consistently affecting agent hyperparameters. Meta-learning encompasses techniques where elicited emotion modifies learning parameters, while in action-selection works it is employed as a main factor of exploration/exploitation. For example, in [195] emotion coefficients modified the agent's learning rate. Distinctly, in [196] coefficients instead tuned action confidence. Yet, there has also been an overlap of the two functionalities, as in [192] the emotion element both served as the learning rate and tuned exploration randomness. Additionally, there is also reward modification, where emotion factors in the reward function as occurred in [197], and state modification, in which emotional values integrate state-space and adjust Q-Values as in [193]. Overall, as overviewed in [23], the majority of emotional RL techniques elicit emotion from stimuli appraisal, representing it dimensionally, and apply it to reward/state modification. Contrarily, value/reward function-based elicitation and meta-learning as a function are the least common categories in this research area. Regardless, choosing elicitation, type, and function modes is highly dependent on problem/task specifics and constitutes a problem in and of itself, considering emotion's debatable nature. More importantly, the positive impact of emotion is self-evident as all works report better learning or behavioral adaptability in some form, stemming directly or indirectly from its integration into artificial agents or robots.

On a separate note, AI adaptability also encompasses plasticity in terms of parameter and structural optimization concerning single-task, multi-task, and other objectives. Still, ANN weight tuning aside, design typically is rigid in that architectural arrangement and parameterization remain static during learning and inference phases [198]. Approaches to this issue have been varied. Some attempt selective activation of network branches or elements, either conditionally or through the use of masking. For instance, [199] grew network depth progressively according to hardware capabilities whilst dropout masking was used for the reduction of network width with similar objectives in [200]. Nonetheless, these fail to adapt dynamically during training and thus provide little benefit to learning. Relatedly, masking has also recently been applied to gradient updates [201], adaptively and congruent with biological functioning in that changes only affect parameters deemed relevant to a current task [202]. While this technique does enable more plastic architectures via sectional or joint network training, it raises concerns over impaired convergence. Data-based topology optimization and dynamic routing are also possibilities in graph neural networks [203], following real neural circuitry. However, adopting this methodology raises a panoply of other issues. While several other approaches exist, techniques that employ some form of neuron activity tracking seem most promising. For example, [204, 205] proposed a cumulative trace of the product between pre- and post-synaptic neuron activity, self-modulated by a network-computed signal, to adjust weights according to connection eligibility. Also, in [206], weight updates rely on a contrastive predictive loss function wherein connection potentiation/depression depends on the accordance of expected and actual neuron activity. Not only are these latter works more akin to real neural processes, but they also provide a great deal of plasticity and flexibility towards data, albeit with increased complexity and computational requirements.

Plasticity of hyperparameters other than structural is likewise explored outside of emotional RL research, given the complexity of manual tuning on increasingly larger or smaller but trickier ANNs [207]. Plus, as the optimization of any network parameter has been deemed possible through minimization concerning variational free energy [208], the recent trend of bio-inspired methodology addressing that same objective appears highly appropriate, given the high likelihood of real neural circuitry adapting and behaving by the free energy principle [209]. To exemplify, current works focusing on adaptive learning rate in ANNs have proposed layer-wise scheduling of the parameter [210] and neuron-wise rate variability consistent with the alignment between forward and feedback neuron activity [211], both congruent with real neural functioning. On a separate note, in [212] activation function parameters were also varied according to cell-specific dynamics. These works all resulted in considerably faster convergence when compared to less biologically plausible approaches, in support of the mentioned hypothesis. Furthermore, other researchers have begun to counter the rigidity of ANN parameterization by introducing signals mimicking neuronal dynamics, which have led to superior performance when compared with conventional approaches [213]. Naturally, this generated discontent for the lack of adaptability in current ANN development frameworks.

The works mentioned in this overview all provided AI some level of increased adaptability, having done so through either direct emulation of neural processes or by abstracting learning/behavioral characteristics that prove advantageous in real life and then integrating these in the current methodology. Nonetheless, this area is only presently becoming mainstream, with many paths of research in neuromorphic AI design still being unexplored. Plus, while we may be distant from the goal of a self-developing and autonomous artificial agent, its precursors are already becoming a very real possibility, as is the case for self-procreation [214] and self-motivated open-world learning [215]. Yet to achieve these and related goals, intrinsic motivation to learn or act and structural/parameter optimization are required capabilities. The former can likely be achieved through the integration of emotion in an agent's cognitive process, while cell or layer activity monitoring is a promising approach for the latter. Consequently, current AI research should prioritize the development of these capabilities in agents.

Chapter 7

Emulating Dopamine

Contents

This chapter explores the effects of emulating the neuromodulator dopamine as a hyperparameter of artificial neurons. This work was originally carried out in an attempt to hasten convergence of DL models towards ideal solutions, yet later moved more towards observing the parallelism between biological and artificial dopamine effects. First, substantial weight variation was employed as signaling of dopamine release and consequential boosting of strength change in neuron connections. Subsequently, this scheme was allied with $D1$ and $D2$ receptor functioning, respectively to observe the impact potentiated excitation and suppressed inhibition would have over model convergence. While obtained results have yet to warrant a publication, this content is still presented considering its emulative interest.

7.1 Context

The integration of neurophysiological characteristics in AI methodology has proved beneficial in the past for greater understanding of brain development [216]. Relatedly, the interpretation of neural network functioning under a neurobiological premise has also provided leeway into how these systems may be further improved [217]. Naturally, there is a growing motivation to consider a range of brain-like phenomena when designing and/or evaluating AI frameworks to advance their efficiency while also providing interdisciplinary insight.

The adaptive convenience of synaptic plasticity, and particularly the potentiation role dopamine has over this process, has constituted a significant motive for its computational modeling [43]. Particularly in terms of AI and deep learning, synaptic plasticity through dopaminergic emulation has been presented as a convergence catalyst in DL [218], among other benefits. This is achieved through the application of the potentiation-silencing dichotomy of connections, typical of the brain, in ANN architectures. Nonetheless, this remains a somewhat overlooked research topic. Some techniques that target it often formulate this neuromodulatory process via Hebbian learning. For instance, in [219] authors considered both modulatory and non-modulatory neurons, so weight updates of non-modulatory connections were mediated by the sum of modulatory inputs to the corresponding downstream neurons and a Hebbian rule. This both improved model performance and prevented catastrophic forgetting. Tracing of neuron activity is also occasionally used to influence neighboring neuron connections. In [220], artificial neurons boast conventional and plastic weights, the latter of which is based on a Hebbian trace retaining recent activity information. A dopamine-like signal is then used to gate this component, improving the fine-tuning of a CNN in a transfer learning scenario. The similar approach of [204] also achieved improved performance and lower perplexity when learning distinct tasks. Other options for neuromodulated plasticity include considering separate networks to adapt activation parameters in the main branch, leading to faster learning as in [221]. While these works demonstrate the potential of introducing neuromodulated plasticity in ANNs, techniques are often not readily applicable to standard DL frameworks in their current form. Instead, authors typically develop ANN mechanics from scratch to test their hypodetheses. While beneficial for customization, this is both time-consuming and often impeding result reproducibility. Moreover, neuromodulatory process emulations developed for neural networks are generally not faithful enough to possibilitate analogy with biological bases, perpetuating the separation between brain-inspired AI and the brain [217]. Naturally, mitigating this issue could aid in bolstering further interdisciplinary collaboration, aside from the benefits bio-inspiration brings to AI and DL research.

This short experiment focused on an emulation of dopaminergic neuron impact over ANN performance, combining aspects of both [219] and [220] yet instead employed over the standard Keras framework [222] for DL. The main goal encompassed assessing whether any improvements would emerge in terms of model convergence speed, as well as discussing any potential similarities with biological counterparts. To achieve this, activity tracing was implemented in scattered positions across a shallow network architecture. Connections between neurons with this ability and regular units were then affected by the trace, with stronger ones being potentiated while weaker ones were further silenced. Initially, the implementation was kept simple, with dopaminergic neurons influencing their regular counterparts directly. Subsequently, $D1$ and $D2$ receptor behaviors were also emulated for comparison. A standard CIFAR10 [223] object recognition task was employed for testing these approaches.

7.2 General Overview

As mentioned, this formulation considered dopaminergic artificial neurons as separate from their conventional neighbors, akin to the reduced and somewhat centralized dopaminergic cell groups that synthesize dopamine in the brain [43]. Thus, these represented a fraction of units within ANN layers, boasting an additional parameter Tr compared to their conventional neighbors. As per equation 7.1, the parameter is designed to accumulate in a dopaminergic neuron i the mean difference in strength of connection with neurons k from the preceding layer, as training progresses to the current iteration t from its antecedent.

$$Tr_i(t) = \gamma \cdot \Sigma_{k=1}^{N} \frac{w_{i,k}(t) - w_{i,k}(t-1)}{N} + (1-\gamma) \cdot Tr_i(t-1) \tag{7.1}$$

The γ factor here determines the impact of new mean differences over Tr. Assuming the trace surpasses a preset threshold, then it may influence the synaptic connections of neighboring neurons, as depicted in Fig. 7.1. The rationale behind this process stems from monitoring a real dopaminergic neuron's excitability and its necessity for a strong enough action potential at the presynaptic terminal to freely diffuse dopamine into the synaptic cleft [224].

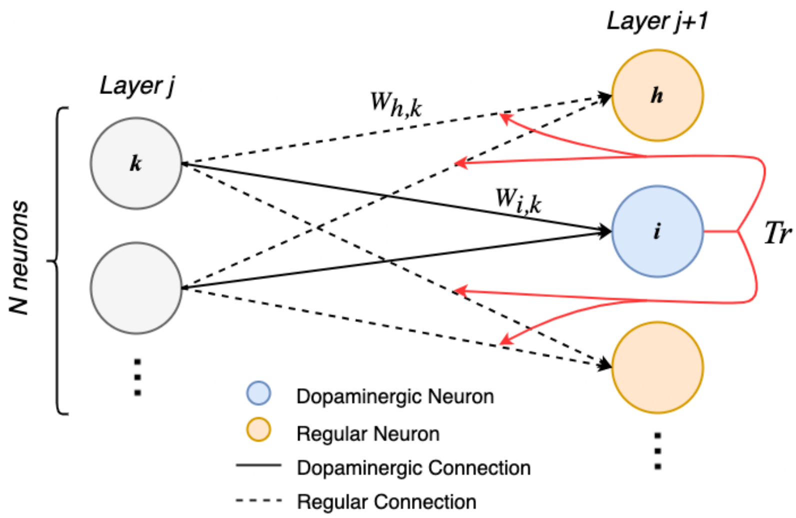


Figure 7.1: Structural overview of a dopaminergic neuron in a fully connected network layer, accumulating a trace Tr based on its connections with neurons from a preceding layer. The same trace is then used to influence connections of neighboring conventional artificial cells.

7.2.1 Basic Influence

In the first instance of implementation, Tr was made to directly potentiate or weaken neighboring synapses. Thus, for each neuron h, the output passed on to neurons in succeeding layers followed equation 7.2. Here, *dop* represents the set of dopaminergic neurons in the current layer, σ is the activation function, and η adjusts the impact of a dopaminergic trace based on the spatial proximity of neuron h to dopaminergic neuron i.

$$x_h(t) = \sigma\{\Sigma_{k=inputs} \left[\underbrace{(w_{h,k} + \Sigma_{i \in dop} \eta_i \cdot Tr_i(t))}_{\text{synaptic strength}} \cdot x_k(t) \right] \}, h \notin dop \tag{7.2}$$

With this formulation, it becomes possible for synaptic connections of conventional neurons to be either potentiated or suppressed based on the traces accumulated by dopaminergic neurons. As traces are indicators of positive and negative change in dopaminergic synapses, they can gauge dopamine release and thus be employed as a direct influence (for the sake of simplicity). This direct influence is meant to mimic the dopamine-dependent shaping of plasticity in a real brain [225], which can boost learning efficiency via the mentioned silencing/amplification of unimportant/important connections. Plus, adaptability towards novelty also improves as new experiences are more easily consolidated by plastic components in real neural structures [226]. In addition to this formulation, while real dopaminergic neurons do boast self-feedback regulatory mechanisms [227], this implementation did not consider such type of inter-neuron influence. Hence, dopaminergic neuron output followed a standard linear combination of its input. On the other hand, conventional neuron output was impacted by the proximity-weighted traces of neighboring dopaminergic cells, forming the *synaptic strength* factor of 7.2.

7.2.2 $D1$ & $D2$ Receptors

Following the initial implementation of dopamine effects over conventional neurons, a second iteration of this framework was developed considering the intervention of $D1$ and $D2$ dopamine receptors. As detailed in section 3.3, $D2$ receptors at postsynaptic cells suppress inhibitory transmission from presynaptic neurons, whereas $D1$ receptors exert an opposing action by extending it further than it would otherwise last and vice-versa for excitatory transmissions. Moreover, considering traces Tr as a simplistic direct dopaminergic influence over neighboring cells, these can be either excitatory or inhibitory to their connection weights depending on whether they are positive or negative, respectively. Thus, conventional cells of this framework can be considered to have additional gating parameters simulating receptor behavior, similar to Fig. 7.2.

Naturally, each conventional neuron would have its own unique $D1$ and $D2$ gating behaviors, these being variants of the general behavior depicted above. For the sake of simplicity, these variations were considered as random parameter differences for each receptor behavior function in the form of step and backward step functions for $D2$ and $D1$ receptors respectively. Adapting eq. 7.2 to include this $D1|D2$ gating simply results in eq. 7.3.

$$x_h(t) = \sigma\{\Sigma_{k=inputs}\left[\left(w_{h,k} + G_{D1|D2}\{\Sigma_{i\in dop}\eta_i \cdot Tr_i(t)\}\right) \cdot x_k(t)\right]\}, h \notin dop \qquad (7.3)$$

7.3 Observations

In order to test these formulations, object recognition was learned by a simple VGG-like architecture considering the CIFAR10 dataset [223]. The architecture itself was composed of two convolutional layers, each followed by a respective max pooling layer. The result of this network segment underwent 50% dropout and was then fed to a fully connected layer for final classification. Within this architecture, the

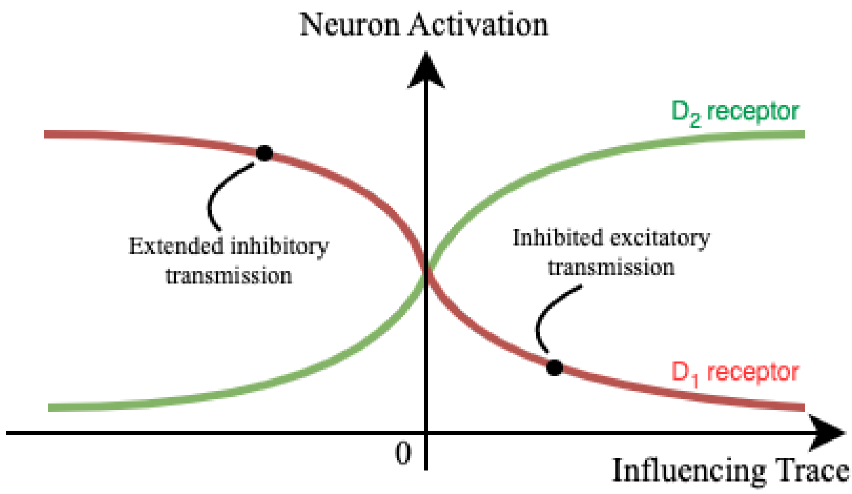


Figure 7.2: Exemplary behavior of $D1$ and $D2$ receptors at post-synaptic conventional neurons, in terms of extension or reduction of excitatory and inhibitory signaling.

developed dopaminergic-like framework was implemented as a new type of fully connected layer wherein several cells take on the dopaminergic behavior described. This new layer was used as a replacement of the classification fully-connected layer in the architecture, for performance comparison. In terms of hyperparameters, the γ factor was kept at 0.3 so the past accumulated change would not be forgotten and remain impactful for the trace, despite new change. The Adam optimizer was employed for model training, considering the default learning rate of 0.001 across all tests which encompassed 100 epochs each with a batch size of 64 and a validation split of 20%. Finally, the proximity parameter η was implemented as decaying according to neuron distance, by having conventional neurons consider η_i as 0.5 for immediately neighboring dopaminergic cells and sequentially halved per each cell location moving sideways on the fully connected layer. This is exemplified in Fig. 7.3.

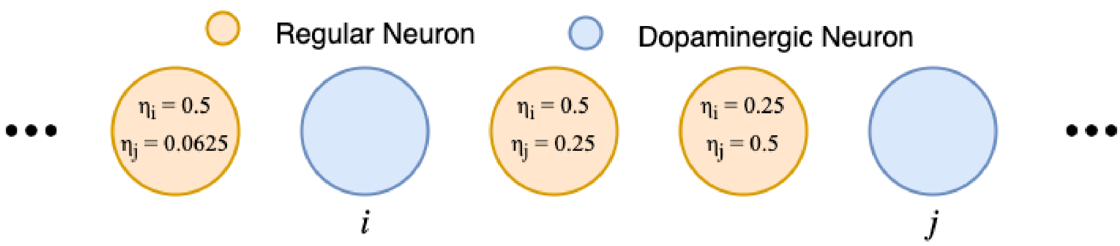


Figure 7.3: Example progression of η halving according to neuron distance in a fully connected layer.

7.3.1 Performance Efficiency

To assess the usefulness of this emulative framework, a comparison was performed between the conventional architecture, the basic dopaminergic, and the $D1/D2$ dopaminergic architectures described to assess whether learning was more efficient with the latter ones. The dopaminergic architectures were varied in terms of the number of dopaminergic neurons, scattered as $10-50\%$ of the cells in the replaced fully-connected layers. Results from this comparison are shown in Fig. 7.4, displaying both accuracy and loss curves for the architectures considered.

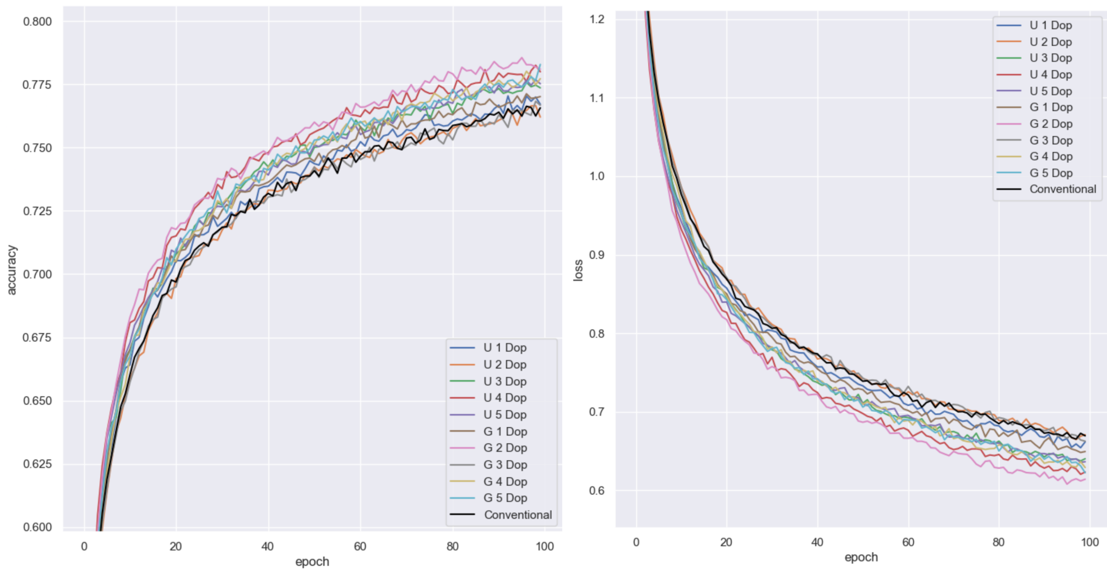


Figure 7.4: Detail of accuracy and loss curves for a conventional architecture, assessed against a dopaminergic architecture, either non-gated (U) or gated (G) and with 1 to 5 dopaminergic neurons among 10 total cells, in terms of performance efficiency improvement.

7.3.2 Adaptability Towards Novelty

The second experiment of this work envisioned understanding how the dopaminergic architectures would handle the introduction of novel data instances during training, compared to their conventional counterpart. Considering the dopaminergic potentiation/depression mechanism, it could be possible for new data to be assimilated faster or for forgetting to occur less in such architectures when compared to conventional ones. Thus, the CIFAR dataset was split into five slices, each containing all examples of 2 random classes, and added to the training dataset of a model periodically over its 100 epochs of learning. Specifically, training started with classes $\{0, 1\}$ at epoch 0, then progressed to classes $\{0, 1, 2, 3\}$ at epoch 20, and so forth until all classes were considered by epoch 80. Again, the dopaminergic architectures were varied in terms of the number of dopaminergic neurons, scattered as $10-50\%$ of the cells in the replaced fully connected layers, while the conventional architecture was kept intact. Results from this comparison are shown in Fig. 7.5, displaying both accuracy and loss curves for the architectures considered.

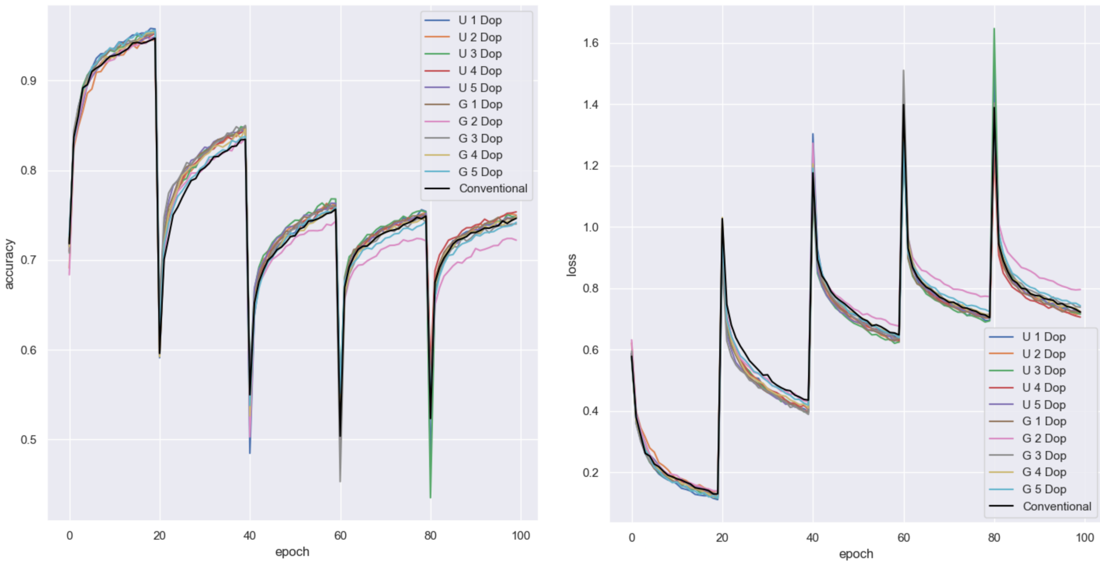


Figure 7.5: Detail of accuracy and loss curves for a conventional architecture, assessed against a dopaminergic architecture, either non-gated (U) or gated (G) and with 1 to 5 dopaminergic neurons among 10 total cells, in terms of adaptability towards data novelty.

7.4 Discussion

Results from the first experiment are depicted by proximity in both accuracy and loss curves across all architectures. Nevertheless, a small increment in terms of accuracy tied with a decrement in loss can be observed when comparing the performance of the conventional architecture against dopaminergic ones. Additionally, a clear separation cannot be made between non-gated and $D1/D2$ gates dopaminergic architectures. For instance, while the best performance has been attained for the gated dopaminergic architecture with 2 non-conventional cells, the same architecture with 5 non-conventional cells does not perform better than a non-gated architecture with 4 dopaminergic neurons. However, it can be observed that the non-gated architecture with 2 dopaminergic neurons and the gated architecture with 3 dopaminergic neurons appear to be the closest in terms of performance to the conventional architecture.

In general, the observed curve proximity precludes a conclusion that dopaminergic architectures provide a significant improvement in terms of learning performance against conventional architectures. While this increase does appear to exist on this preliminary testing, standard backpropagation training procedures may be impeding it from being greater by suppressing the impact of dopaminergic traces in a manner akin to what occurs to noise. Regardless, result variability with different numbers of dopaminergic cells may suggest there is an optimal quantity to be found for either non-gated or gated architectures, also dependent on the total number of neurons. Further testing with larger and more heterogeneous architectures would be required to verify this hypothesis.

In terms of the experiment regarding adaptability towards data novelty, spikes were to be expected and can be observed in the accuracy and loss curves, corresponding to the introduction of novel classes in the training dataset. This is also naturally accompanied by an initial overall decrease in accuracy and a rise in loss. Additionally, proximity in both accuracy and loss curves can again be observed across all architectures. Unlike the first test, however, the separation between the conventional and dopaminergic architectures appears even slimmer, with no single architecture maintaining a lead performance throughout the whole training period. This is clear from how different architectures spike at different moments of class introduction. For example, a gated dopaminergic architecture with 3 non-conventional cells does so at the introduction of classes $\{6, 7\}$ at epoch 60, yet at epoch 80 with the introduction of classes $\{8, 9\}$ it is the same but non-gated architecture which spikes the most.

The shortage of separation between performance curves again supports a lack of usability for this dopaminergic framework. Nevertheless, it could be possible that the suitability of an architecture towards the introduction of particular data instances, while already considering other specific data, is also dependent on its dopaminergic characteristics. To test for this more robustly, different data class combinations and modes of introduction during training should be considered along with the architectural variation. Moreover, other learning tasks should be evaluated with various ANN types in order to generalize such conclusions.

Overall it can be concluded that results from this emulation, either non-gated or gated, were rather underwhelming. Withal, observations were interesting and constituted a first step towards a dopaminergic formulation for ANNs, since they provided valuable insight into its requirements and usefulness. Certain aspects should be considered for the future, however, such as spatial correlations in data whose importance could be better-captured thanks to the implemented dopaminergic spatial impact mechanism. While on fully-connected layers this is not as noteworthy given their longitudinal arrangement, architectures such as CNNs could benefit from such a characteristic considering their specificity to spatial data analysis. Hence, while the obtained results were not ideal, it is not necessarily true that this same concept could not be further scrutinized and become useful across other tasks and deep learning methodologies.

Chapter 8

Exploration from Internal Drives

Contents

In this chapter a phenomenon through which epistemic and achievement emotions mediate knowledge exploration is examined for AI, inspired by observations in humans reported in cognitive psychology studies. The goal was to enable artificial agents to analyze data under their perceived needs, emulating human learning autonomy. Thus, based on a neurophysiological background, RL and SL methodologies were combined to develop a model capable of learning the mentioned emotion-exploration relationship whilst performing a generic task. Results corroborated psychological findings in humans and were presented in [228], from which this chapter's content is partially adapted. Additionally, a hypothesis was put forth regarding the impact of simultaneous emotional states over exploration, to be explored further in cognitive psychology. Finally, it was demonstrated how artificial emotion can be objectively useful for AI as an exploratory drive and autonomy catalyst, given the learning advantages it entails.

8.1 Context

Emotion, as a natural phenomenon of sentient life, has an impact far beyond its social characteristics. As detailed in Chapters 3 and 4, it intervenes in a panoply of neural processes involved with learning, environmental salience, and its exploration, which allow humans and other animals to better navigate through their day-to-day activities. Whilst endowing artificial agents with real emotion appears dubious still, developing the conditions with the current methodology for byproducts of emotion to manifest as well as be advantageous for their tasks is a real possibility. Regardless, studies that explore this are virtually nonexistent despite the advantages to learning autonomy being self-evident and obtained results also representing a suitable way to corroborate Psychology and/or Neurophysiology. Relatedly, exploration is a fundamental aspect of cognitive development and independent behavior in human beings [229]. Specifically, the onset of confirmation bias influences the way information is actively sought to ratify prior beliefs and inference [230]. Thus, if AI is ever to grow autonomous, research must strive to develop similar behaviors.

As mentioned, epistemic and achievement states, respectively emotions which are triggered by cognitive incongruity and intrinsic success [79] such as surprise and pride, are mediators of knowledge acquisition and learning given the triggering or halting effect they have on exploratory behavior. These traits are both currently lacking and highly desirable for autonomous AI methodology. Specifically, state-of-the-art surveying has posited the need for an exploratory drive to oversee new data acquisition for long-term autonomy to become a possibility [231]. Otherwise, without it an agent is limited to the small set of data it knows, possessing no control over its quality or repeatability. Issues such as overfitting and lack of generalizability are natural consequences. Hence this problem ties with the trade-off limitations of AI in terms of its exploration-exploitation dichotomy and the open challenge that is designing systems whose learning is supported by intrinsic motivations [232]. Nonetheless, recent reviews argue these limitations may be mitigated via a synergy with RL components [214]. While few areas have approached autonomous exploration in AI via RL and/or some interpretation of internal motivation, such as robot path planning [233] and navigation [189], published techniques consistently target their domains only. Moreover, these are typically restrictive and rely on optimizing the order by which all data is examined, regardless of usefulness, or on following a preset escape protocol when unplanned scenarios are encountered. Contrarily, the presented study aimed to understand if an internal exploratory drive could be sourced from artificial emotion, akin to real life, which entails some scarcity of closely related literature. Nevertheless, that literature could include [234], where a latent dynamics model was endowed with Bayesian surprise as the dissimilarity from its posterior to prior beliefs to reward exploration, or [235] whose process estimates the change in prediction error (PE) as a metric of learning progress so if low, exploration is shifted to more interesting change-inducing goals. Both reported efficient and effective exploration. Also in [236], analogously to pride effects, authors suggest a form of intrinsic rewarding reliant on competence progress, based on which agents can explore their goal space.

Assessing the effects of emotional states such as surprise and pride on exploration is not exclusively beneficial to AI research. As advocated in [237], ANNs constitute a valid framework on which to explore a range of phenomena covered by Psychology and Neuroscience. Thus, should the obtained results match the findings of these areas, they can serve as additional corroboration or go even further by proposing additional paths to research in those fields. This process is not unlike what has been presented in [238], where ANN testing showed responses were following probability matching, when in a RL scenario. This further supports the suitability of AI models to emulate animal behaviors typically observed when studying the matching law. For these reasons, one way to formulate corroborative experiments is to establish a hypothesis based on findings reported in Psychology or Neuroscientific studies, devise a biologically inspired artificial system that follows the respectively derived theory or theories, and apply the final product to scenarios analogous with real-life learning. As a consequence, the resulting performance may be relevant for interdisciplinary comparison. Currently, this procedure is not intentionally followed by any mainstream research, to the best of our knowledge.

In terms of the proposed experiment, it stems from the Psychological research of Vogl *et al.* on the origins and outcomes of epistemic and achievement emotion [79, 80]. Specifically, a group of within-person studies on task-related emotional outcome and interest assessment demonstrated correlations between high-confidence errors, the onset of epistemic and achievement states, and their precipitation of curiosity and exploration. The (here abridged) results demonstrated that:

- $R1$ - Surprise produced by cognitive incongruity positively predicts curiosity, which itself leads to an increase in exploratory behavior.

- $R2$ - Pride stemming from intrinsic success positively predicts exploratory behavior, though not necessarily after correct answers.

Stemming from these observations, our approach was to design an analogous experimental scenario using DL methodology and determine whether the same conclusions could be drawn from artificial learning. To this end, participants were emulated as artificial agents, themselves combinations of task-oriented and RL systems. For each of these, a task-oriented system was subjected to a SL testing procedure while being mediated by feedback from the RL system. Naturally, our architecture closely followed the theory and neurophysiologic circuitry seemingly most responsible for the findings reported in [79].

8.2 System Background & Design

The link between the increase of attention activity towards certain features and encoding of feature-specific PE error (rPE), as observed in frontostriatal circuits, strongly resonates with the direct relationship between epistemic/achievement emotion and the onset of PE or internally attributed reward [50], which has been extensively reported in cognitive psychology. In particular, authors have denoted dopaminergic neurons of the VTA as mediators of "liking" responsiveness to stimuli [239] and encoders of rPE, shifting their firing to feature-specific stimuli post-learning experience [240]. Concurrently, the VS projects particularly strong surprise signals for goal-relevant features, contributing an attentional bias towards relevant matter and consequential high reward attainment. Cortical areas have likewise been observed to track positive affect and pride, with activation largely targeting the VS when task outcomes depend on performance decisions [241]. Moreover, SN neurons are known to engage with attentional functions and project to the DS, which itself

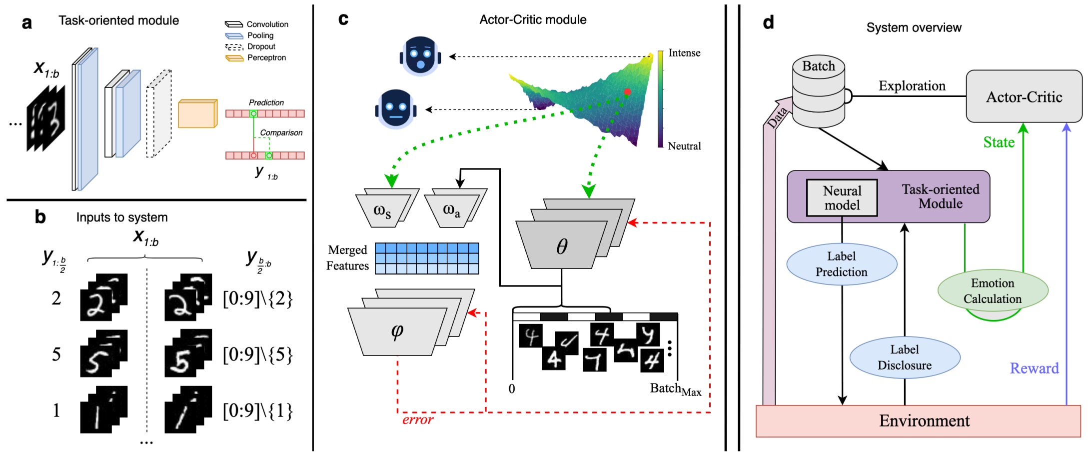


Figure 8.1: The proposed system employs a task-oriented module and a RL actor-critic module to associate emotion and exploration in a way conducive to improved performance in a given task. **a**, The task-oriented module first samples one data instance from the environment, to perform a simple classification task. It does this via a pre-trained neural model, whose convolutional layers extract meaningful visual information. **b**, The loaded data encompasses handwritten digit images from a dataset partially adulterated so that half of its labels will not match with their respective instances' visual content. **c**, The actor-critic module is composed of two separate neural models, for the actor and the critic respectively. The variable accuracy resulting from the task-oriented model is compounded with a random high-confidence score, to compute an epistemic or achievement emotion, according to reports in cognitive psychology research. The actor model θ receives this emotional score (either of pride or surprise) as its sole input and decides on an appropriate exploratory rate for the task-oriented model. The critic model also receives a computed emotional score as input to its branch ω_s, in addition to the actor's chosen exploration rate on its ω_a branch. The resulting merged features are processed by ϕ to generate a feedback signal scrutinizing the actor's decision and the critic's performance. **d**, The AI system performs this routine continuously, sampling a new instance whose task-oriented evaluation triggers an emotional response, then processed into the actor-chosen exploratory rate. In turn, this determines the size of a same-type data batch to be analyzed in the following step.

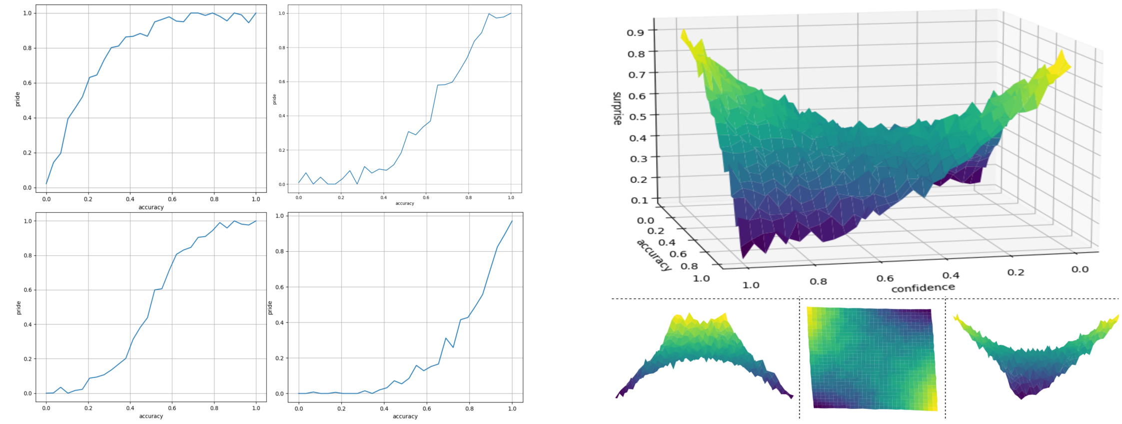
has been demonstrated as paramount to the successful expression of surprise-induced learning functions, yet not so much on their establishment [242]. These observations firmly indicate the involvement of basal ganglia (BG) circuitry, and principally its classic actor-critic structuring, in modulating feature attention and learning parameters. Plus, modulation of focus and knowledge exploration by the BG is strongly supported by literature [243, 244, 245, 246]. Following this paradigm, the critic, commonly associated with the VS, employs its rPE-based attentional shifting and performance/outcome appraisal capacity to adjust the reactive processes put into play by the actor, represented by the DS. Various other neural structures involved in the current task will also be influenced as well as relay information back to this BG circuitry, supporting the exploration-exploitation dichotomy.

Based on this neurophysiologic foundation of the BG, an artificial agent architecture was designed wherein an actor-critic module (Fig. 8.1C) sustains the causality relationship between epistemic/achievement emotions and the exploration rate of a task-oriented module (Fig. 8.1A). This design drew from RL methodology to account for the neuron signaling process which enables self-appraisal and increase of attention towards stimuli inducive of PE. Moreover, the system was endowed with epistemic and achievement emotion formulae so this signaling would stem directly from the feelings of surprise and pride. These formulae were fixed and factor in metrics of the task-oriented module while implicated in a cognitive task. Finally, to complete the action-reaction loop, this latter model's activity was made dependent on the rate generated by the BG-like architecture.

Considering knowledge exploration within a guided task is a form of directed rather than stochastic exploration, a deterministic approach was hypothesized to fit this behavior more adequately. Thus, DDPGs were deemed appropriate and implemented as the actor-critic module emulative of the biological BG circuitry. Within an artificial agent, both the actor and critic receive as input the current state of that agent, which is made up of its solo or combined emotions. With this information, the actor outputs the exploration rate to be passed on to the task-oriented module, while the critic judges the actor's decision indirectly based on the task-oriented module's performance with the provided parameter. A simple SL task was selected for this latter module as any more complex tasks would be out-of-scope for this work. Hence, it interacts with the environment (i.e. the experimental scenario) by making predictions over input data and being informed of their veracity. Results of this disclosure constitute the task outcome, based on which error is computed, implicitly following the mechanics of other limbic circuitry. Reward is also provided by the environment, depending on task performance. These signals are then used to optimize the actor and critic models. This overview is depicted in Fig. 8.1D.

8.2.1 Emotion Functions

As mentioned, cultivating real emotion in AI is still a dubious claim. However, we maintain that replicating cognitive conditions promotive of epistemic and achievement emotion is achievable within AI by considering performance metrics or other scores as condition determinants. For instance, accuracy may serve as a pointer of error and achievement considering it gauges model correctness on a task. As a consequence, accuracy spikes may be interpreted as increasing success, whereas de-escalation entails a less favorable scenario. Pride variations could thus trail behind task accuracy variations, corresponding to personal achievement or lack thereof [247]. This accuracy-pride matching would likely entail a curve of positive slope and unknown convexity with small variations (see Fig. 8.2A). The set of representable emotions may be broadened by factoring in additional pointers, such as confidence, besides standard performance metrics. For example, high-confidence errors trigger the feeling of surprise as a result of their inherent cognitive incongruity. Plus, surprise may also be induced from insecure or unexpected attainment of success [248], while reduction of this feeling may occur in each of these scenarios if confidence is to decrease or increase, respectively. Accordingly, surprise appears to boast a saddle-like behavior, with polarized variations of accuracy and confidence implying intense bursts of this feeling, whilst matching magnitudes of the two factors indicate reduced or emotional lack thereof (see Fig. 8.2B). These perspectives on surprise and pride are widely backed by cognitive psychology literature [79, 80, 249, 250], supporting their explicit implementation as drivers of AI behavior.



(A) Pride (B) Surprise

Figure 8.2: Example curves following a positive prediction of pride based on increasing accuracy (left) and a multiple perspective surface view demonstrating how the emotion of surprise may correlate with accuracy and confidence (right), both of which stem from cognitive psychology research [79, 80].

Considering Fig. 8.2, multiple alternatives may represent surprise and pride, and thus a single set of functions, meeting the requirements defined by psychology research, was selected at random for the experimental part of this work. These factored in performance metrics of the task-oriented module, with accuracy a being readily available and already bound to the $[0,1]$ interval. Therefore, a positive prediction of pride P based on task accuracy could be obtained from (8.1), which was chosen for this achievement emotion.

$$\begin{aligned} &Pride\colon [0,1] \rightarrow [0,1] \\ &\qquad a \mapsto Clip\left[(100 \cdot C_1)^{-(a-1)^2} + \mathcal{N}(\mu, \sigma^2)\right],\ C_1 > 1 \end{aligned} \tag{8.1}$$

This function considers $C_1 > 1$ and added Gaussian noise $\mathcal{N}$ to account for any variability related with personality differences. Overall clipping secures the bounding of the emotion to an acceptable range of $[0,1]$. On a separate note, surprise computation additionally considered a confidence score c bound to the interval $[0.8,1]$. This was introduced to simulate the high levels of confidence expected of the task-oriented module, whose training ensured top performance. The saddle-like rough surface of this epistemic emotion S could therefore be obtained from (8.2), which was chosen for the experiments.

$$\begin{aligned} &Surprise\colon [0,1] \times [0,1] \rightarrow [0,1] \\ &\qquad c, a \mapsto Clip\left[\mathcal{T}\left(\mathcal{R}\left(a^2 - c^2\right)\right) + 0.5 + \mathcal{N}(\mu, \sigma^2)\right] \end{aligned} \tag{8.2}$$

Here, $\mathcal{R}$ denotes a $45° \pm C_2$ rotation around the surface's saddle point, with $C_2 \in [-20°, 20°]$. This is followed by a translation $\mathcal{T}$ to each domain interval midway (i.e. $[0.5, 0.5]$), making it so low-confidence success and high-confidence mistakes are met with high surprise, while opposite scenarios induce less or no surprise. Similarly to pride, Gaussian noise is introduced for added variability and clipping ensures bounding to the $[0, 1]$ interval. Finally, $\mathcal{N}(0, 0.03)$ was employed for both functions, along with combinations of randomly generated C_1 and C_2 values. The goal of this was to obtain varied artificial agents with individual differences yet following the same grand pattern, much like human participants.

8.2.2 Overview of Experimental Scenario

This proposal attempted to mimic the experiments of [79, 80], wherein a human participant was informed of 20 potentially incorrect general knowledge statements, one by one, and asked to determine their veracity. Immediately after each, the participant was informed of their accuracy and evaluated on various epistemic and achievement emotion scales. Additionally, the participant had an exploratory option to request additional statements on the same topic which caused cognitive incongruity, demonstrating interest and curiosity. This process was repeated for a large number of participants over 2 studies, to correlate the several factors evaluated.

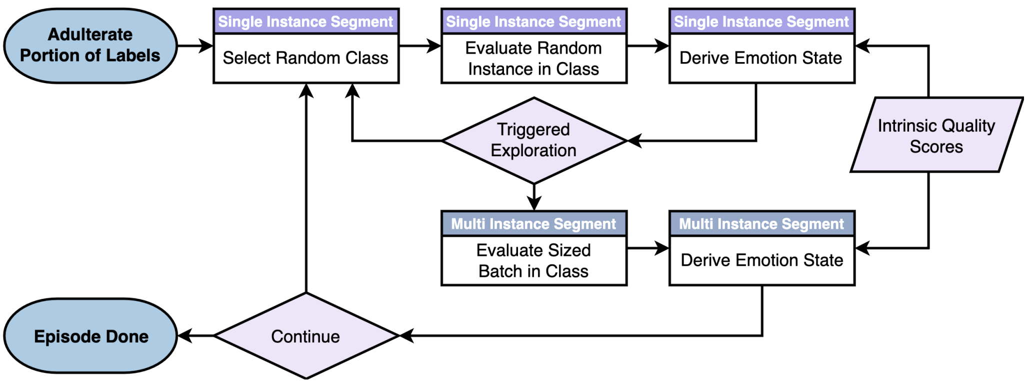


Figure 8.3: Flow diagram of a RL episode, which composes the experimental scenario designed to replicate in artificial agents the same psychological testing procedure used by Vogl *et al.* with human participants [79, 80].

In our artificial emulation, the task-oriented module employed a model to classify handwritten digit images from the MNIST dataset [251]. This model was pre-trained on half the dataset, achieving a test accuracy above 99% and near 0 loss. The unused half had a percentage of its labels adulterated (see Fig. 8.1B) to simulate the original studies' statements with varying degrees of truth and prompt high-confidence errors in an agent when classified by its pre-trained task-oriented model. Thus, an agent is first provided with a single random instance from the subset with adulterated labels. The classification of this instance and outcome disclosure (based on label correctness) induces an emotional state in the artificial agent. This happens accordingly with the functions described above, using single instance accuracy and the random yet fixed high confidence value. This fixation is purposeful, to represent the tonic nature of confidence in real life. The DDPG section of the agent then uses the derived state to decide how much the task-oriented system will be exploring ($\mu_{exploration}$) in the coming segment of the step, with a set batch maximum of 64. Emotion is likewise derived in the multi-instance segment for later analysis, following the same procedure. A flow diagram of this process is shown in Fig. 8.3.

Given a standard assumption that participants intend to perform well in this activity, artificial agents are given a basis reward whose polarity corresponds to that of the difference between explored batch accuracy a_m and single instance accuracy a_s. This deems exploration useful only if it yields improvement in terms of task performance. A sparse component is also added to the reward, matching the variation of epistemic/achievement emotion that occurs during a step. This serves to either minimize surprise or maximize pride, in an attempt to comply with the free-energy principle which illustrates the necessity of self-organizing agents to reduce uncertainty in future outcomes [252]. In the cases of no exploration, emotion variation is null and reward is obtained directly from the only available accuracy score a_s, normalized to the $[0, 1]$ interval. Considering the $sign function$, this results in:

$$R = \begin{cases} 0.5 \cdot sign(a_m - a_s) \pm \Delta emotion & , 0 < \mu_{exploration} \leq 1 \\ 2 \cdot a_s - 1 & , \mu_{exploration} = 0 \end{cases} \tag{8.3}$$

Finally, the evaluation of 20 single instances and their subsequent batch analyses constitutes an episode, following Vogl's 20 statement procedure, with each agent performing a fixed number of episodes. Expectantly, analysis of a single instance happening to have an incorrect label, following the described process and considering a well-trained SL model, results in a high-confidence error (low accuracy - high confidence). In theory, the agent should learn to associate the consequent high surprise with an appropriate exploratory variation targeting the type of data that caused its emotional state. Contrarily, analyzing a legitimately labeled instance should directly yield a high reward without the need for further exploration. A similar but simpler principle applies to pride so that agents should learn to associate its variation with exploration, so higher rewards can be achieved. Moreover, by allowing the agents to decide on their rate of exploration for the data types observed, their actions then dictate surprise minimization and pride maximization in accordance with the free-energy principle. The sparse component of the reward function further backs this objective, by making the emotion variation negative for surprise and positive for pride.

8.3 Results

In [79], two studies were carried out totaling 247 participants. Proportionately, a total of 250 distinct artificial agents were created by introducing distinct random variations in the free parameters of their emotion function, and initial confidence scores. Each of the artificial agents was made to undergo the experimental procedure for 100 training episodes, with obtained results being averaged across the total number of participants in the experiment. Moreover, the usage of a multiple-class dataset was intended to draw general conclusions from results more invariant to any effects of data class/type which may occur with data of single or low-class amount.

8.3.1 Surprise/Pride vs Exploration

Towards assessing the correlation between knowledge exploration and epistemic or achievement states, surprise and pride functions were first implemented separately in the experimental process. Results of this experiment are shown in Fig. 8.4. These paralleled Vogl's studies as similar conclusions were obtained for statements $R1$ and $R2$.

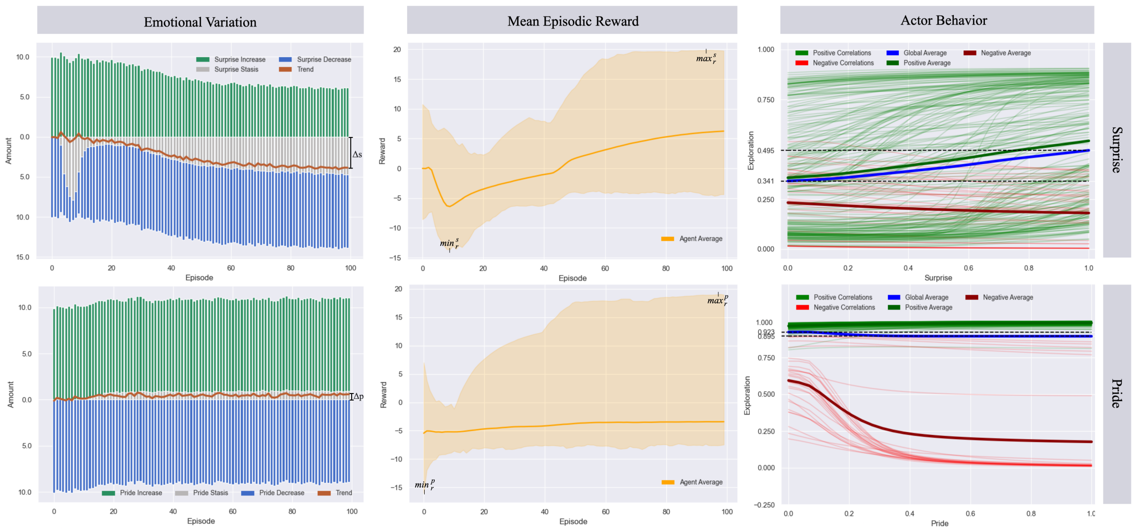


Figure 8.4: Results for surprise and pride as separate exploratory drives. Leftmost column: Episodic mean of emotion differential between single sample and subsequent batch analysis steps, across all implemented agents over the entire learning cycle. Middle column: Mean cumulative reward obtained by agents at each episode of the cycle. Rightmost column: Mean actor behavior at the end of the learning cycle, correlating surprise/pride with exploration.

Cumulative reward increased over learning episodes for either emotion (Fig. 8.4 middle column) and plateaued by the end of the cycle. This indicates model convergence and ensures the behaviors learned by the artificial agents are not random. In terms of surprise (top), reward peaked at $max_r^s = 19.87$, rising from a minimum $min_r^s = -13.64$ and averaging at 6.26 by the end of the cycle. As for pride (bottom), reward fluctuated between a minimum $min_r^p = -15.97$ and a maximum $max_r^p = 19.05$, whilst its final value averaged at -3.41. The mean trend of cumulative reward also demonstrates an increase in both emotions. However, pride boasts only a slight growth whereas surprise exhibits a short depression in earlier episodes followed by a steady increase later on. Regardless, both indicate agents successfully learned to match emotional states to actions and consequentially improved their performance. The success of the experiment can also be corroborated by the fluctuation of emotion observed in artificial agents over time (Fig. 8.4 first column). As can be seen, the initial variation is well-balanced for both pride and surprise as the amount of increases matches that of decreases in the first learning episode. Over time, surprise appears to reduce or stagnate as bursts of this emotion become on average $\Delta s = 38.52\%$ less frequent by the final episode, even despite the clear preference for stasis in the first 10 episodes. Though not as prominent, pride average variation still shows an upward tendency, with decreases occurring $\Delta p = 5.90\%$ fewer times by the end of the cycle. Increases also do not appear to be favored over stasis for pride, as the observed trend is mainly fostered by the latter.

After the experiment and despite random differences and noise, agents demonstrated similar and monotonic behaviors relating exploration with either emotion (Fig. 8.4 third column). A causation effect was most evident for the surprise experiment, which resulted in agents displaying a 15.4% increase in exploration stemming from greater surprise. This positive correlation was observed in 217 of all agents, outshining the remainder of 33 who displayed a slightly negative correlation. As for pride, a mean deflating effect was instead observed from the full set of agents, obtained from 222 positive weak correlations and 28 negative correlations. Out of these, 22 correlations amply decreased exploration between 25% and 75%, which is reflected by the modest 2.8% exploratory reduction observed for increasing pride.

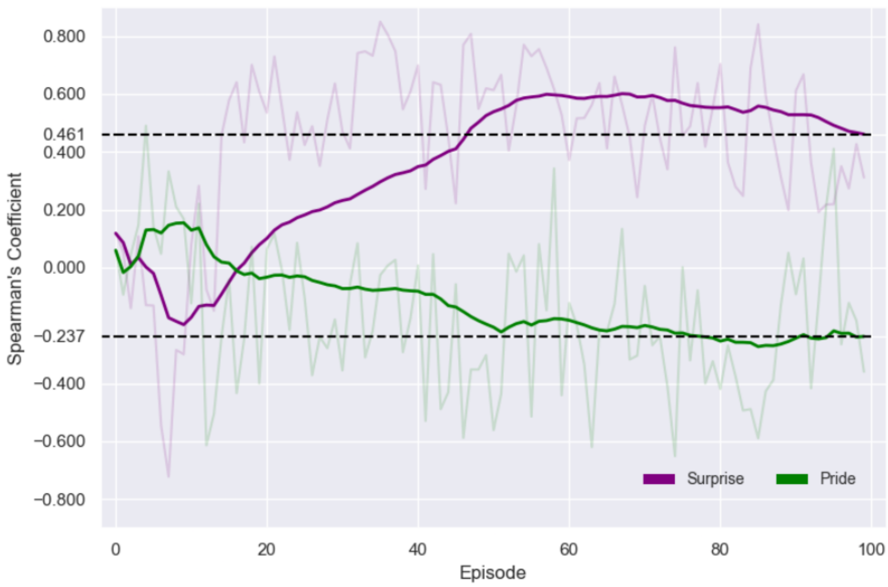


Figure 8.5: Agent episodic mean of Spearman's correlation coefficient between actor-chosen exploratory rate and its causal surprise or pride score (pale), smoothed by a moving window of 40 samples (bold).

Robustness of the relationships between emotion and exploration also varied throughout the cycle, as shown in Fig. 8.5 by the evolution of Spearman's correlation coefficient ρ [253]. A sliding window of 40 episodes was applied to smoothen the considerable variability observed. These mean trends demonstrate an increase for surprise which results in $\rho_{surprise} = 0.461$, indicating a strong positive correlation with exploration, and a decrease for pride which results in $\rho_{pride} = -0.237$, implying instead a weak negative correlation.

8.3.2 Combo Proposal

In this experiment, surprise and pride functions were implemented together in artificial agents. Results of this experiment are shown in Fig. 8.6, wherein a three-way relationship is demonstrated between exploration, epistemic, and achievement emotion. This is therefore proposed as a topic for future Psychological research to corroborate.

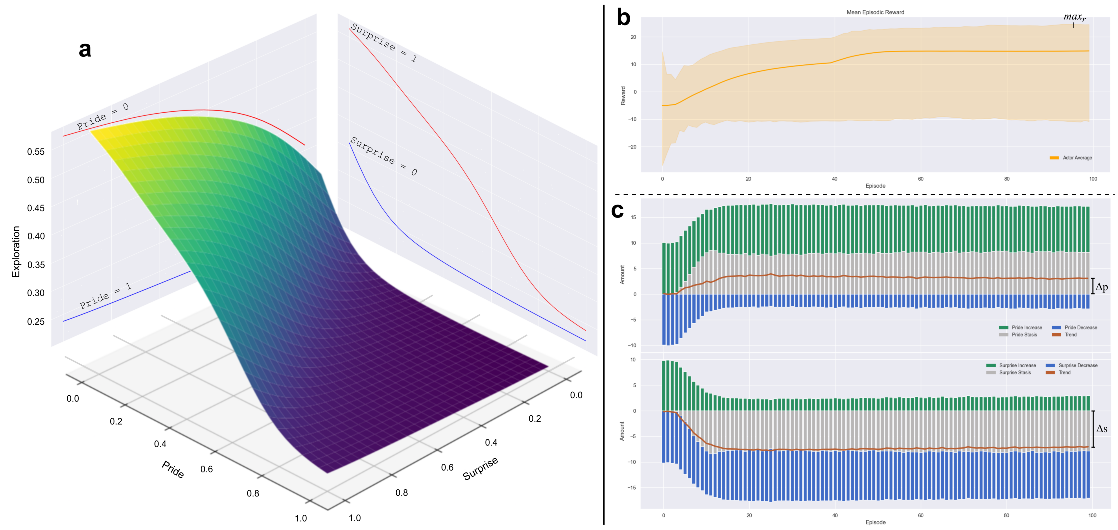


Figure 8.6: Results for surprise and pride combined as one exploratory drive. **a**, Mean actor behavior at the end of the learning cycle, correlating both emotions with exploratory behavior. **b**, Mean cumulative reward obtained by agents at each episode of the cycle. **c**, Episodic mean of emotion differential between a single sample and subsequent batch analysis steps, across all implemented agents over the entire learning cycle.

Compared to the previous experiments, cumulative reward (Fig. 8.6B) here increased much earlier and plateaued at a higher value, with few agents even reaching $max_r = 24.5$ while the average trend settled at 14.86. Thus, the behavior learned by artificial agents is expected to be even more robust than in either single-emotion scenario. Moreover, it is clear how emotional fluctuation (Fig. 8.6C) stabilized quite early and more prominently than before, for both emotions. While initial variation was well-balanced, on average surprise bursts were reduced by $\Delta_s = 70.4\%$ while pride reduction became $\Delta_p = 30.9\%$ less frequent by the last learning episode. However, rather than respective increases/decreases, a preference for stasis is the evident cause for this stabilization. In terms of actor behavior (Fig. 8.6A), the overall trend of exploratory variability appears to be retained for surprise and pride yet at different rates. For instance, while exploration still increases monotonically with surprise, this is largely dampened by the feeling of pride. As the latter reaches its maximum, the relationship between surprise and exploration becomes nearly negligible. Contrarily, pride's sole dampening effect over exploration appears less effective if surprise is also occurring, as agents report higher rates following greater scores of that epistemic emotion. This results in a wave-like surface, peaking exploration at maximum surprise combined with minimum pride levels and smoothly reducing the rate as surprise nears zero and pride spikes.

8.4 Discussion

Results from the separate emotion experiments strongly resonate with emotion-mediated exploration reported for humans. Vogl *et al.* [79] twice demonstrated a causal relationship between surprise and exploration of knowledge in separate studies, evidenced by positive path coefficients obtained for surprise-induced curiosity, and ensuing exploration. Coefficients of 0.285 or 0.262 correlated surprise and exploratory behavior, respectively in the first and second studies. The mean exploratory increase of 15.4% attained by our artificial agents with growing surprise follows the postulated path relationships. Moreover, this parallelism demonstrates how the adoption of human-like behavior by AI grants it a better standing over time when placed under similar testing conditions. Finally, the non-windowed mean Spearman's correlation coefficient of 0.311 we obtained for the final episode indicates the same relationship robustness as the within-person correlation values of either study.

In terms of pride, Vogl's results across studies contrasted with those of surprise by lacking consistency. While both reported negative correlation coefficients, respectively of -0.073 and -0.177, their closeness to zero indicates a weak relationship between pride and exploration, if any. In addition, the second study's negative path coefficient contradicted the first study's positive coefficient, despite both nearing zero. The weak correlation between exploration and this achievement emotion was corroborated by our experiment wherein Spearman's coefficient was likewise negative yet took longer to deviate from null, compared to the surprise experiment, and still stagnated at a lower absolute value. The obtained 2.8% mean exploratory decrease over growing pride is also congruent with the path coefficient of the second study, supporting a dampening effect from this emotion. Nevertheless, this percentage decrease is slim, and considering it was obtained from various strongly divergent agent behaviors, the polar discrepancy of either study's path coefficient appears more valid and demonstrative of the negligible effect pride alone has over exploratory behavior. The fact that pride as a unique positive state may be damaging to cognitive performance [254], for which exploratory behavior is key [255], could justify the observed decrement of exploration. Nonetheless, additional experimentation would be required to assert this hypothesis.

Still, regarding the solo emotion experiments, minimization of surprise was considerably successful, ensuring agent behavior complied with the free-energy principle. The same cannot be concluded for pride, as maximization was feeble. Regardless, both experiments corroborated findings from cognitive psychology, with agents learning to self-mediate knowledge exploration and improve task performance by exploiting internal emotional drives.

The merging of emotional states employed as an exploratory drive, while not drawing inspiration from a psychological research basis, revealed an interesting cross-effect between surprise and pride. Respectively, boosting of minimization and maximization of these emotions over time is clear, as both achieved considerably larger differentials between the first and last learning episodes. Pride increases specifically went from negligible to a steady change, while surprise decreases also peaked at a higher value and stabilized earlier in the experiment. This suggests either emotion has a complementary effect over the other, as their combination led to agents more closely adhering to the free-energy principle. This complementarity is further corroborated by the fact that agents were able to accumulate more reward overall and also significantly sooner in the experiment, compared to the previous solo emotion runs. As for the correlation observed between the emotional mix and exploratory behavior, pride's dampening effect also appears more intense here, whereas before it was only slight. The near elimination of surprise exploration boosting in the presence of pride spikes also supports this notion. Therefore, we would conclude that pride's impact on exploration is indirect and more likely perceptible in the presence of surprise, whose impact is comparably more overt but susceptible to the variations of that achievement emotion.

Overall, this work demonstrated how psychology and neurophysiology combined may provide a basis for research into AI autonomy. Artificial agents performing this sort of emulation can plausibly develop traits useful for learning, as demonstrated by our results. Likewise, theory on human cognition and behavioral traits may be corroborated or scrutinized if AI displays some postulated behavior under similar conditions [256]. Nevertheless, considering agents are likely to remain unconscious and refrain from supporting humans in any foreseeable future[257], it is unlikely they will replace biological participants in experimental scenarios. Thus, while biological emulation is an inspiring topic, comparison of outcomes obtained via AI and human/animal observation should be approached with care.

8.5 Applications

In order to demonstrate the usability of the developed models, two applications have since been designed in the field of social agents and robotics. Both works employed the methodology described in this chapter, more concisely yet still conducive to authentic surprise-exploration correlations as shown in Fig. 8.7. Each resulted in a publication [258, 259], from where the contents of this section are adapted.

8.5.1 Adaptive Attention

This work [258] presented an agent that performed adaptive broadening of its familiarity with user facial features in real time. This functionality reiterated the usability of emotion as a learning/behavioral metaparameter as well as the importance of this chapter's work for considering biological traits when designing artificial agents.

Specifically, a loop was implemented wherein faces within a camera's field of view were detected and a similarity comparison was carried out against knowledge stored in memory already. Each mean comparison value was obtained as a squared Euclidean distance between OpenFace [260] features of a detected face and those of faces known previously. Should a mean of these comparisons surpass a preset threshold, then the face was considered new and its embedding stored in memory.

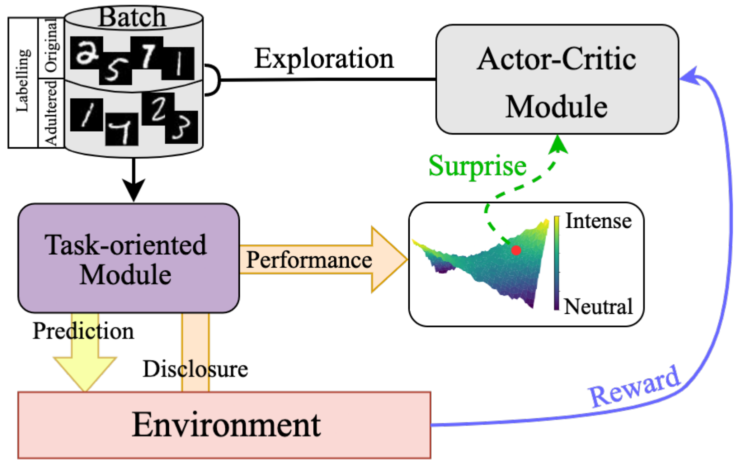


Figure 8.7: Simplified learning loop for a task-agnostic agent to manifest a realistic matching between surprise and exploratory behavior. Task performance is interpreted in terms of surprise induction, which an actor model then uses to infer an adequate exploration ratio. The actor's decision is optimized by a critic model receiving a reward signal from the environment.

Moreover, the mean comparison value of each detected face was interpreted as a surprise score and normalized to the interval $[0, 1]$, with the highest of all being fed to the actor model. This determined how much this most surprising face should be explored, based on which a robotic agent's eyes would stare at the corresponding person. Finally, the purpose of this mediated agent focus was so that the representation of a new, surprising face would be stored in memory and continuously updated until it no longer triggered surprise. A diagram of this process is shown in Fig. 8.8.

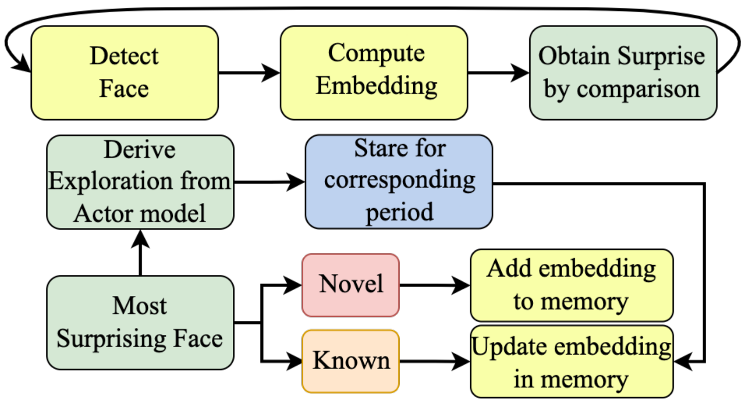


Figure 8.8: Adaptive attention based on facial feature similarity and surprise-induced exploration.

8.5.2 Adaptive Persistence

The experiment presented in [259] envisioned adapting the level of persistence applied by a social robot to its actions, based on surprise-induced exploration. Likewise, to the previous work, this design validates the usability of emotion as a learning/behavioral meta-parameter.

In summary, a Q-learning based decision-making network architecture was implemented for selecting the most appropriate social action at each given interaction state. Each selected action was then adapted by a separate model meant to modulate its persistence, as per a generated exploratory ratio. This latter value stemmed from the artificial emotion of surprise interpreted as the incongruity between an expected and the corresponding detected user reaction, and fed to an actor model trained according to the process described earlier. This matching was likewise employed for the calculation of the reward necessary for the Q-learning algorithm. Moreover, the user reaction was observed within the subsequent interaction state caused by the robot's selected action, forming a continuous loop of actions with persistence optimization. A flow diagram of this process is shown in Fig. 8.9 for easier understanding. The framework was implemented in a NAO robot from Softbank Robotics [261] and an experiment with 22 real users was carried out as depicted in Fig. 8.10.

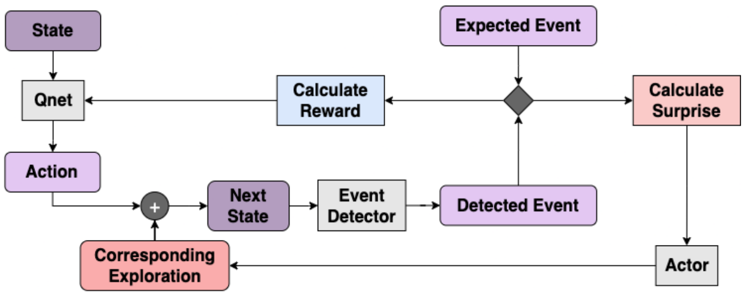


Figure 8.9: Social action and corresponding persistence optimization loop.

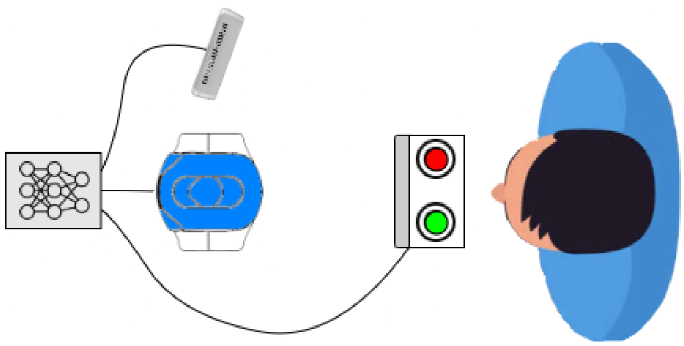


Figure 8.10: Top view of the experimental setup.

Part IV

Final Remarks and Future Work

Prologue

While this thesis has presented a broad perspective on emotional competence and adaptability, it is virtually impossible to develop work in each of their many potential research sub-topics. Moreover, each work presented here is but a fraction of what can be done in its respective sub-topic. As a consequence, future work plans may always entail delving deeper into an existing proposal or devising a new one altogether. Both options have value as the former progresses a field further, while the latter may become a starting point for other members of the research community to develop their works. Nevertheless, given the under-researched stage that bio-like and autonomous AI is currently in, providing new paths to be explored here appears more promising for the future. For this reason, even though future work overviews are provided for both emotional competence and adaptability, another chapter is first included. This chapter is a demonstration of how other sub-topics about this thesis are available and can be explored in the future, apart from the ones presented already in former chapters. Specifically, it presents an idealization of artificial dreaming for ANNs to prevent overfitting, inspired by theories attempting to explain dreaming in biological beings.

Chapter 9
Artificial Dreaming

Contents

The work presented in this chapter drew inspiration from a theoretical viewpoint on biological dreaming as an overfitting coping mechanism. That viewpoint was used as a basis for developing an algorithm for AI to deal with the same problem autonomously, via periodical data augmentation and interpretation based on the model's current state of knowledge. To this end, a data augmentation procedure was developed, which integrates a core version of Google's DeepDream generative system [262], followed by a labeling process contingent on minimizing the impact of generated data over an ANN's state, similar to the method described in the works of Sucholutsky *et al.* [263]. Early testing appears promising and led to the publication of a position paper [264], from which part of this chapter's content is adapted. Regardless, further substantial experimentation is warranted so more solid conclusions may be drawn on the efficacy/efficiency of the proposed method.

9.1 Context

The overfitting issue commonly observed in machine learning applications represents a major setback to general AI and the achievement of greater learning autonomy. More often than not, provisioning of quality data is limited. Thus, as most models lack any form of independent exploration, methods such as network reduction, data augmentation according to pre-set rules, and early stopping are typically implemented to prevent further loss of generalization [265]. Naturally, these require a hands-on approach from users, choosing which network sections to prune, augmentation rules, and stopping criteria based on personal expertise or trial and error. This may not always yield a solution or it may prove too complex for some models, beckoning a necessity for overfitting to be dealt with by the model itself.

Dreaming is one of the most characteristic aspects of sleep. These are essentially virtual concoctions of memory, emotion, and knowledge, recent or consolidated, which result in multi-sensorial experiences and hold disputed significance. Research has postulated an evolutionary origin for dreaming, being developed by the brain as a mechanism to simulate threatening or unsettled matters to determine the course of action most likely to result in survival and success [266, 267]. Reports of unconscious solution perception being catalyzed in dreaming participants during sleep studies [268] further support the idea that dreaming provides new perspectives and information analysis. Moreover, evaluation of internalized problems during sleep is congruent with attempted task integration in dreams [269]. Dream absurdity could also be associated with the emotional factor of subconscious simulations, contributing to a mental preparation by hyperbolization of potential scenarios [270]. Nevertheless, dreamless sleep has likewise been correlated with performance improvement and learning [271]. In the case of non-dream sleep stages, neural pattern replay is critical for abstracting core knowledge and consolidating memory, as evidence suggests [5]. However, the latter is unable to explain the purpose of dreams during other sleep stages. Plus, the odd and consistently scattered nature of dreams disfavors the objective usefulness in the daily life of these unconscious experiences, as noted by [272]. Overall, these observations strongly point to dreaming as a coping mechanism for the homogeneity of data and stimuli experienced during daily life. This is especially true considering how the scarcity of distinct examples perceived during a day is contradictory to a brain's generalization capabilities. Finally, the intuitiveness of the human brain is what enables the incorporation of scattered priors in condensed experiences to generate new intelligence [273]. Such information has recently led to the development of the overfitted brain hypothesis [272]. This theory, strongly backed by neuroscience, posits that the brain dreams to warp statistically proximal instances observed throughout a day and consequently can prevent its overfitting or increase its generalization capabilities. The stochastic corruption of typical sensory input

Sleep is commonly divided into two major types throughout a single session. These are designated as REM (rapid eye movement) when idiosyncratic dreams most often occur, and NREM (non-REM) or thought-like sleep. Clear differences in terms of brain wave activity make for an easy distinction between the two types [274], with slow-wave activity being characteristic of NREM. Here, posterior and central brain regions occasionally present off-states that hinder experience generation [275], and neural pattern replay occurs [5]. As such, events are often recalled as mundane and memory-related. As off-states subside and the brain cycles to REM, it shifts to higher frequency (gamma) activity with more vivid events and a surrealistic tendency [276]. Thus, it is presumable that NREM sleep accounts for the consolidation necessities of the brain, whereas REM deals with the cognitive and data-augmenting aspects detailed by the overfitted brain hypothesis.

Emulation of REM dreaming in neural networks could cause the resulting data warping to function as a source of creativity for AI. In this scenario, the remaining stages of NREM and wakefulness could be considered homologous to conventional training and post-training model usage. While ambitious, this is not the first instance of work correlating sleep phenomena and DL processing. For instance, statistical corruption of data as that performed by Google's DeepDream [262] has been used for multi-candidate dreamed object classification, by correlating real image feature vectors with decoded brain activity patterns obtained from sleeping subjects [277]. Thus, employing these data warping techniques in generating augmented data for ANNs, based on their knowledge state, to prevent overfitting would be both plausible in terms of its biological foundation, and represent a next step for research into AI autonomy.

9.2 Idealization

Adhering to the overfitted brain hypothesis as inspiration, this work focused on developing a form of artificial dreaming wherein data can be augmented or inferred by models mid-training, in an attempt to reduce or prevent the impact of overfitting. Ensuring the process could be integrated into most existing ANN frameworks was also taken into consideration during the design process.

9.2.1 Artificial REM

Brain activity during REM resembles that of its awake state, with seemingly random bursts of intense, fast, and out-of-sync waves. Moreover, the body enters a separate homeostatic balance wherein optimized functioning is no longer its target [278]. Data is then processed differently, as the brain forgoes risk aversion and considers novel associations. Plausibly, information flowing through the brain at a given moment has patterns drawn from general knowledge, memory, and episodic information [279] forced on it to simulate possible future scenarios. These patterns may already correspond to content similar to the initial information or, as the dream progresses, become increasingly decipherable as novel albeit unconventional data instances. Figure 9.1 depicts this hypothesis using visual data, despite its applicability to other modalities.

This theory can be adjusted for AI. For instance, initial data may be noise from random activation of certain neural regions or external stimuli to a dormant brain, which is directly re-creatable in standard DL methodology. It may also stem from out-of-context memory accesses induced by the shifted electrical connectivity in the brain. This lack of context can be achieved in neural networks by simply presenting them with other data, unrelated to their task. Homologously to the statistical warping of information performed by a dreaming brain, ANNs can be made to force global patterns of learned data on those initial inputs through excitation of one or more of the layers processing them (i.e. DeepDream's procedure) [262]. An example of this augmentation is shown in Figure 9.2.

After securing self-fabricated data examples, ANN models must then interpret and understand their content, similarly to how the brain attempts to make sense of dreams as they occur. Naturally, labels assigned to these augmented instances must reflect both original characteristics and those integrated by maximization of layer activations, entailing a strong level of ambiguity. Regardless, this combination of distinct traits in a shared data instance is what may enable new knowledge inference and disrupt overfitting in the model. Moreover, the ambiguity of that augmented data is also congruent with real dream recalling (often fuzzy). Such a level of imprecision may be accounted for by employing soft rather than hard labels, ensuring

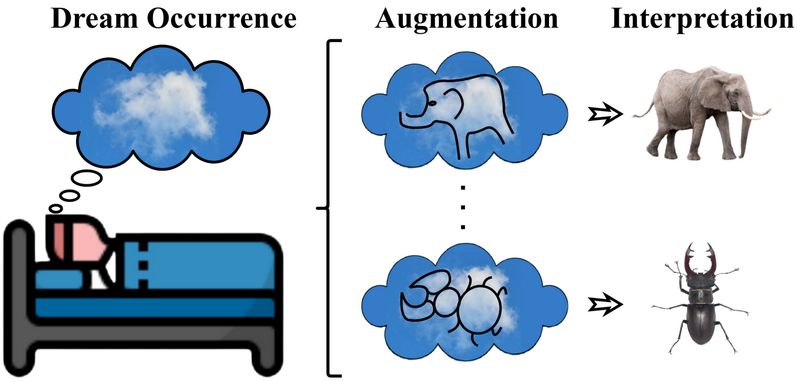


Figure 9.1: Dream data augmentation, from left to right. The dream starts either by random memory access or activation of brain structures. Based on current knowledge, personal interest, persisting thoughts, and others, the dream is morphed to match the current state of the brain. Simultaneously it attempts to interpret this information in order to form a response to it. As a side effect, this helps negate latent overfitting.

Figure 9.2: Deepdream's maximization of layer activations used for augmentation of a rocky formation image (left), results in buildings resembling pagodas (right). This was done using a network trained on places by MIT Computer Science and AI Laboratory as presented in [262]. The new image could be useful for disrupting overfitting, as the network continues training on place identification, using soft labels to account for its augmented characteristics (i.e. the pagodas).

data usability post-dreaming. These soft labels can be made learnable as dream data is iteratively interpreted according to its respective model's current knowledge. Thus, the labels would be optimized, to reduce the impacts of their instances over the model's regular training and prevent catastrophic forgetting from hindering its task performance. This part of the proposal is inspired by Soft-Label Dataset Distillation [263], where a similar notion has been employed successfully. In it, labels are perfected so as to minimize the error of a model over real data, when trained with a single forward pass of distilled data.

Finally, in case dream-generated data is found to be nonsensical as the network is unable to find an interpretation for it, this procedure requires an escape option. Humans typically deal with dream absurdity by disregarding the corresponding episodes completely. Respectively, a similar process may be implemented for ANNs to assign an absurdity class to augmented data deemed meaningless. Instances with an acceptable interpretation could then be added to the network's current dataset and increase heterogeneity during training, whereas nonsense data would simply be discarded.

9.2.2 Algorithm Design

This algorithm relies on typical ANN nomenclature to remain generalizable to most existing frameworks. For the same reason, it was designed using the same basic notation as [263] and related works. This presumes a K-layered neural network f parameterized by θ, with typical backpropagation based on a twice-differentiable loss function $l_1(x_i^r, y_i^r, \theta)$. The goal is to find an optimal set of parameters θ^*, using a training dataset $\mathbf{r} = \{x_i^r, y_i^r\}_{i=1}^N$, according to (9.1).

$$\theta_{new} = arg\,\underset{\theta}{min} \frac{1}{N} \sum_{i=1}^{N} l_1(x_i^r, y_i^r, \theta) \triangleq arg\,\underset{\theta}{min}\, l_1(\boldsymbol{x}^r, \boldsymbol{y}^r, \theta) \tag{9.1}$$

This optimization is performed iteratively with stochastic gradient descent being computed against a batch of training data using (9.2), with a learning rate η. Additional parameters could be considered, such as momentum α, yet were disregarded in this algorithm for the sake of simplicity.

$$\theta_{t+1} = \theta_t - \eta \nabla_{\theta_t} l_1(\boldsymbol{x}_{batch}^r, \boldsymbol{y}_{batch}^r, \theta_t) \tag{9.2}$$

Assuming that a network shows early signs of overfitting at a moment $t > 1$ of training, then this process should be halted and the model switched to dreaming mode. This mode can be divided into 4 major phases, as shown in Fig. 9.3, according to our idealization: initialization, augmentation, interpretation, and assessment. The first phase of initialization parallels the random flow of data characteristic of REM sleep, represented by a dream dataset $\mathbf{d} = \{x_i^d, y_i^d\}_{i=1}^L$. Instances of this dataset, henceforth named augmentation *themes*, make up the general setting of a model's dream and will have patterns forced on them according to its current state of knowledge. Moreover, these instances should only encompass content dissimilar from that represented in the regular dataset $\mathbf{r}$ used for training. Accordingly, dataset $\mathbf{d}$ may be created in several ways, depending on data availability and application goals. Pure noise instances can be used, in which case dream *themes* have no inherent meaning. Contrastingly, real data may be sourced from another dataset, provided it is unrelated to the training dataset yet available in its same modality. To exemplify, face images may be used as *themes* for a model being trained for object recognition with the CIFAR10 [280] datasets.

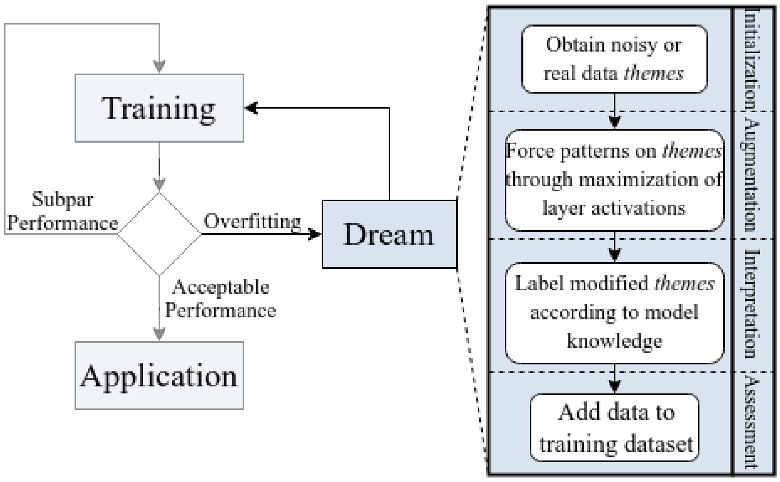


Figure 9.3: Overview of the neural network's workflow, with the dream stage being used as a mechanism to tackle overfitting.

The augmentation phase is aimed at incepting network knowledge over the dream *themes* of dataset **d**, according to DeepDream's procedure [262]. This augmentation is achieved through the excitation of a few randomly chosen layers $k \in K \setminus \{input, output\}$, which will force over the *themes* any patterns they typically enhance during regular functioning. This presupposes a loss objective $\mathcal{L}_2$ which depends on layer activations $a = f_\theta(\boldsymbol{x}^d)$, meaning the chosen layers' outputs given a forward pass of a *theme*. Since different abstraction and feature levels are dealt with by separate layers, the depth and extraction purpose of each chosen layer will determine the characteristics of the patterns to be imposed over *theme* instances. Hence, the objective $\mathcal{L}_2$ can be defined with a differentiable loss function $l_2(xd_{batch}, a_k)$ based on layer activation according to (9.3), to be calculated individually for each *theme* using (9.4).

$$\mathcal{L}_2(\tilde{\boldsymbol{x}}^d, \theta_t) := l_2(\tilde{\boldsymbol{x}}^d, a_k) \tag{9.3}$$

$$\tilde{\boldsymbol{x}}^d_{new} = \underset{\tilde{\boldsymbol{x}}^d}{arg\,max}\, \mathcal{L}_2(\tilde{\boldsymbol{x}}^d, \theta_t) = \underset{\tilde{\boldsymbol{x}}^d}{arg\,max}\, l_2(\tilde{\boldsymbol{x}}^d, a_k) \tag{9.4}$$

Interpretation of the new, potentially meaningful information now incepted on *themes* constitutes a dream's third phase. The goal here is to optimize the labels for the augmented data to enable the minimization of a loss objective $\mathcal{L}_1$, corresponding to the differentiable loss function typically used in backpropagation as shown in (9.5), in a single step $\theta_1 = \theta_0 - \eta \nabla_{\theta_0} l_1(\tilde{\boldsymbol{x}}^d, \tilde{\boldsymbol{y}}^d, \theta_0)$.

$$\mathcal{L}_1(\tilde{\boldsymbol{x}}^d, \tilde{\boldsymbol{y}}^d; \theta_0) := l_1(\boldsymbol{x}^r, \boldsymbol{y}^r, \theta_1) \tag{9.5}$$

Unlike soft-label dataset distillation [263], only $\tilde{\boldsymbol{y}}^d$ are considered here, whilst $\tilde{\boldsymbol{x}}^d$ are disregarded for this minimization. That is because the model should merely interpret the augmented dream *themes* rather than optimize them as well, according to its current state of knowledge. Consequently, label updates should adhere to equation (9.6).

$$\tilde{\boldsymbol{y}}^d_{new} = \underset{\tilde{\boldsymbol{y}}^d}{arg\,min}\, \mathcal{L}_1(\tilde{\boldsymbol{x}}^d, \tilde{\boldsymbol{y}}^d; \theta_0) = \underset{\tilde{\boldsymbol{y}}^d}{arg\,min}\, l_1(\boldsymbol{x}^r, \boldsymbol{y}^r, \theta_0 - \eta\nabla_{\theta_0} l_1(\tilde{\boldsymbol{x}}^d, \tilde{\boldsymbol{y}}^d, \theta_0)) \tag{9.6}$$

By performing the second and third phases of this artificial dreaming procedure several times, not only do forced patterns become more prevalent on each dream *theme* but their labels become increasingly more reliable. Furthermore, the number of classes represented by these labels may be extended to include an additional class pertaining to absurdity or non-meaningful data. Naturally, instances majorly assigned to this class should be dropped by the algorithm when entering its final phase: assessment. Here, the augmented data instances paired with their optimized soft labels are integrated into the regular training dataset **r**, and regular training is resumed. Intuitively according to the overfitted brain hypothesis, this should help disrupt overfitting as the neural model analyses these new data samples in tandem with the task-related data. Algorithm 1 realizes the overview provided in this section.

Algorithm 1 REM Dream Emulation for Overfitting Disruption

Input: M: Number of themes to occur during dream; α: step size; n: batch size; T: Dream depth in steps; $\tilde{y}^d_0$: initial value for $\tilde{y}^d$.
Output: Augmented dream data ($\tilde{\boldsymbol{x}}^d$,$\tilde{\boldsymbol{y}}^d$).

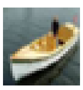
Dream Data of Theme Initialisation :

- $\tilde{\boldsymbol{y}}^d = \{\tilde{y}^d_i\}^M_{i=1} \leftarrow \tilde{y}^d_0$
- $\tilde{\boldsymbol{x}}^d = \{\tilde{x}^d_i\}^M_{i=1}$ randomly OR sample batch from dream dataset **d**
- **for** each training step $t = 1$ to T **do**
  - **for** each layer $k \in \{$randomly chosen layers$\}$ **do**
    - **for** dream theme $\tilde{x}^d_i$ **do**
      - Forward pass the theme
      - Evaluate objective function on activations $\mathcal{L}^{(k,i)}_2 = l_2(\tilde{x}^d_i, a_k)$
    - **end for**
    - Compute updated model parameters with SGD: $\theta_1 = \theta_0 - \eta\nabla_{\theta_0} l_1(\tilde{x}^d, \tilde{y}^d, \theta_0)$
    - Evaluate objective function on real training data: $\mathcal{L}^{(k)}_1 = l_1(\boldsymbol{x}^r_{batch}, \boldsymbol{y}^r_{batch}, \theta^{(k)}_1)$
  - **end for**
  - Update dream data: $\tilde{\boldsymbol{y}}^d \leftarrow \tilde{\boldsymbol{y}}^d - \alpha\nabla_{\tilde{\boldsymbol{y}}^d} \sum_k \mathcal{L}^{(k)}_1$, and $\tilde{x}^d_i \leftarrow \tilde{x}^d_i + \alpha\nabla_{\tilde{x}^d_i} \sum_k \mathcal{L}^{(k,i)}_2$
- **end for**

9.3 Preliminary Results

While this bio-inspired algorithm is a working idea, its potential can be observed using a simple example as that demonstrated in Table 9.1. In this case, handwritten digit recognition was considered as the main task being learned by a shallow CNN with two convolutional layers using the MNIST dataset [251]. CIFAR10 [280] images are evidently unrelated to this task and thus may serve as dream *themes* on which to force patterns. This was done on two randomly chosen CIFAR10 images by exciting either the first or the second convolutional layers separately. Even without performing soft-label optimization, a single forward pass of these augmented images through the model reveals that these hold meaningful handwritten digit information. For instance, the image of a ship enhanced by the second convolutional layer of the original model produces something interpreted by the network as resembling digits 2, 3, and 5. Moreover, considering this CNN was not yet overfitted, low loss supports the validity of these predictions. Additional digits or partial features could potentially be further interpreted by the model, were it overfitted, or if more augmentation iterations were carried out. With optimized labels reflecting this content multitude, these instances could be useful for disrupting the homogeneity of the original training data.

Table 9.1: Exemplary run of two CIFAR10 images as themes (top - airplane, bottom - ship) over a single dream iteration, using a double-layered CNN trained exclusively for MNIST handwritten digit recognition. 'Original' rows show the untouched CIFAR10 images, while 'Conv1' and 'Conv2' each refer to an 800-step run of the Deepdream technique over the CIFAR10 images activating the first and second convolutional layers, respectively. Probabilities, shown as percentages, refer to the evaluation of the resulting images by the MNIST-trained CNN (i.e. the soft labels they would attribute after this initial iteration).

		0	1	2	3	4	5	6	7	8	9
Original		13.7	4.8	15.9	6.5	14.8	5.2	10.2	2.3	23.6	3.0
Conv1		0.0	0.0	10.1	0.0	0.0	0.0	0.0	0.0	89.9	0.0
Conv2		0.0	0.0	83.1	1.3	0.1	12.1	0.0	0.7	2.7	0.0
Original		41.4	0.3	16.8	5.0	3.7	0.6	16.2	0.0	15.8	0.2
Conv1		0.0	0.0	98.9	0.9	0.0	0.0	0.0	0.0	0.2	0.0
Conv2		0.0	0.0	66.0	7.8	0.0	26.1	0.0	0.1	0.0	0.0

9.4 Discussion

This work sought to develop a way for the effects of overfitting to be reduced autonomously in AI methodology. Following the brain's ability to deal with this issue via sleep and dreaming, an attempt was made at emulating the latter in ANNs. This involved halting the regular training of a model, followed by an augmentation step wherein patterns detected by its layers were forced over data unrelated to training. The resulting information was then interpreted according to model knowledge, with labels optimized to have little impact over model parameters, and integrated with regular training data to reduce its homogeneity.

While this technique can produce data holding meaningful content to the conventional training, it is largely distinct from common generative methodology. First, architecture agnosticism entails its applicability to most neural network designs. Moreover, it does not require additional models or network branches to be integrated into the main architecture or perform augmentation-specific training. In comparison with such techniques, ours does not incur this additional memory usage and computational costs. An interpretation phase according to model knowledge is also inexistent in typical generative methodology, with labels often being transcribed from related real data or inferred from latent space distribution.

Naturally, there are limitations to be considered when implementing our algorithm. For instance, gradient ascent is used to corrupt dream *themes* with patterns retained by model layers. Similarly to gradient descent, this increases the computational cost of conventional training, with deeper layers requiring further steps for their activations to be maximized over the *themes*. Nonetheless, this is coherent with the brain's heightened activity during REM sleep which matches and may even exceed that of wakefulness, as demonstrated by energy expenditure. Despite this, deeper layers are still preferable for augmentation as those are more likely to incept meaningful information and prevent reliance on the absurdity class of the soft labels. Otherwise, shallower dreams are less likely interpretable by the network and resulting instances may be overwhelmingly disregarded, in which case the dreaming process loses its usefulness in terms of mitigating the effects of overfitting. Overall, these issues can potentially be alleviated by increasing dream depth. Thus, provided there are no critical time constraints imposed for training, enough steps can be carried out for them to become negligible.

In terms of validation, it is also necessary to design experimental scenarios to assess the efficacy of the proposed algorithm against overfitting. These should consider factors such as architecture depth, as efficacy may be hindered with shallower networks. Different data sources must also be evaluated for dream *themes*, to determine if some are indeed more adequate depending on the task being learned by a model. Likewise, it would be interesting to explore what influences layer suitability for excitation during a dream. This could greatly reduce the time consumed by the model as, instead of choosing random layers until dream augmented data is considered useful, appropriate instances could promptly be generated and passed on to the interpretation phase. Finally, if the algorithm does prove effective against overfitting, it will further support the usefulness of neuroscientific theory and findings for addressing common issues in AI and improve its adaptability.

Chapter 10

Final Remarks

Contents

This thesis explored topics related to bio-inspired artificial intelligence, both from a socio-interactive application side based on emotional competence, as well as from a learning adaptability standpoint with neurophysiological bases. This chapter summarizes primary contributions and presents potential paths for future research.

10.1 Conclusion

This thesis covered two major themes related to bio-inspired artificial intelligence, namely emotional competence, and adaptability. The main goal of this research was to further the autonomy of AI systems within the context of interaction and learning, benefiting from insight into naturally optimized processes observed in real beings. The pursuit of this goal was partly motivated by a desire to mitigate societal problems on the growing need for assistance, which is unmatched by current personnel standards, social companionship, and lack of interactivity among people. While emotional competence is mostly trivial for humans and a major adjuvant for the appropriate attainment of these requirements in daily life, its artificial recreation is far from straightforward. Therefore it was imperative to address two main areas of artificial emotion in this document. These were its expressions, encompassing an understanding of user perception and behavioral appropriateness by artificial agents, and its recognizability by those same agents, this latter being a gateway to the fluidity we expect of peers during an interaction.

The fuzzification of the emotional features approach, which was first proposed as an extension of previous work on recognition, constituted the first contribution of this thesis. While human emotional expression is often archetypal in that there is one predominant state being manifested at a time, interpreting it is not a straightforward task. First, states are not necessarily overt, and a lack of characteristic cues may impose constraints or require analysis to inefficiently ponder over the possibilities of a categorical approach through exclusion. Considering information gathered by a DL approach may not necessarily pertain to the desired class, embedding fuzzification can help to still benefit from that information. This ensures the current state is examined based on its correlation with more apparent cues and precludes the necessity for further user sampling, which may be bothersome and have detrimental effects on perception. The proposed approach was capable of performing this process within a discrete group of emotions expressed in prosodic speech features, where recognition is often miscalculated. Hence it was guaranteed that both speaker-specific cues and inter-emotional correlations were considered for improved accuracy, justifying its integration in user-adaptive social behavior algorithms for artificial agents. Validation of the approach was carried out in a controlled manner through the use of emotional speech databases which constitute the current experimental standard in the area. A performance improvement was observed when compared to the non-fuzzy baseline architecture developed in previous works. Furthermore, the basis for the approach is not necessarily unique to the speech modality, as the same principle holds for other forms of emotional expression.

Pertaining to an artificial expression of emotion, a RL-based methodology was presented for the development of empathic behavior in social agents. Nonverbal communication, which empathy integrates, is a major aspect of interpersonal relationships. Its absence is consequently found to be disconcerting and may hinder the acceptance of artificial systems originally incapable of evolving emotional skills. Assistive and companionship services are often faced with situational delicacy wherein the need for the latter is exacerbated. Here, user information ranging from physical cues to deliberate speech is already gathered and updated constantly according to service demands. Hence, techniques such as the one proposed may be integrated into those systems almost seamlessly to form a repertoire of more lifelike social skills. Respectively, facial analysis and explicit feedback were leveraged to recognize a user's emotional state and via reinforcement adapt a robotic agent's facial configuration to match that of the user's. Validation was carried out with real users during a conversation scenario, with posterior surveying corroborating the expectation that empathy would improve HRI. Given the generality of the experimental setup, conclusions are likewise applicable to assistive and social services.

Contributions to greater autonomy and less reliance on human aid aligned more with the adaptability section of this thesis. This is because there is motivation to improve current solutions in social, but more importantly industrial technology, to capacitate it for conformance and flexibility towards the variety of issues that may occur in a work environment. Natural bio-optimization has made it so living beings developed this ability for survival, thus making their neural features the best source of inspiration for adaptable AI. Consequently, adaptability was addressed at three major levels, first encompassing cell-based learning modulation, and secondly emulating neural structure interactivity. The former constituted an attempt at improving learning efficiency as well as providing a better understanding of neuromodulatory effects, whereas the latter targeted natural behavioral development by benefiting from interdisciplinary knowledge. The third level of adaptability considered here focused on the re-creation of a brain-wide mechanism for coping with overfitting autonomously.

While the mimicking of neuromodulatory processes served as a major source of inspiration for other bio-inspired works in this thesis, the results of this particular trial were underwhelming. The emulation of a dopaminergic boosting to neural connections could entail benefits for artificial neurons similar to those shared by their biological counterparts, through the potentiation of meaningful synapses paralleled by suppression of unimportant ones. The goal of this formulation was to improve learning efficiency, reducing the time needed for training as well as the amount of data. A straightforward approach to achieve this goal was to create an activity trace between two neurons, similar to the signaling of dopamine release, which in turn would influence the synapse between the two cells. Regardless, the effect of this weight change acceleration appeared to be insignificant, possibly due to the dopaminergic trace being considered as noise by standard backpropagation training procedures. An additional emulation of D1 and D2 dopaminergic receptors in the artificial neurons was considered as a next step, to observe whether these would improve the impact of dopamine release compared to the basic activity trace approach. While improvement was not significant, it was interesting to observe the overall behavior of this neuromodulator emulation and consider possible paths for future development.

The emulation of neural structure interactivity arguably constituted the greatest contribution of this thesis, both towards AI as well as interdisciplinary research. This is because the association of psychological findings with a neurophysiological foundation for the deep learning methodology presented in Chapter 8 not only furthered AI autonomy but also provided a framework on which to test human behavior hypotheses artificially. The generality retained in the technique enables further studies on the relationship between various emotional states and possible human behaviors they influence to be studied empirically. Consequently, this may adopted by researchers in various fields, constituting a contribution greater than other works presented here. Specifically, experimental conditions subjected to human participants during a psychological study were adapted for application to artificial RL agents implementing an emulation of the neurophysiological circuitry responsible for associating emotional activation with behavioral optimization in living beings. Replication over hundreds of artificial agents demonstrated that the correlations learned, addressing the emotions of surprise and pride respectively over their influence towards exploratory behavior, matched those demonstrated by the human participants in the basic psychological study. The applicability range of these models is wide in AI, enabling researchers to increase the independence level of their systems. It entails an autonomous exploration estimate based on artificial emotion, which can be adapted according to application specifics much like the spin-off works of this topic, developed for proof of concept in the field of social robotics. Additionally, it was further demonstrated that when combined, the emotions of surprise and pride can have a stabilizing effect over each of their respective impacts on exploratory behavior. This was proposed as a research topic for Psychology to validate, conversely to the process followed in this work.

The final work included in this thesis was presented as an open-ended research topic. A theory on the dreaming mechanism employed by the brain to cope with overfitting was considered for ANNs, in an attempt to attribute these the same benefits boasted by their biological counterparts. Adaptability towards overfitting can boost the autonomy of continual learning systems by reducing the necessity for solutions such as user-added data variability. Moreover, datatype agnosticism entails broader applicability in the real world, as industrial services typically require a level of flexibility and issue compliance which most other solutions are unable to cope with. In technical terms, the proposed methodology encompassed forcing patterns of knowledge acquired by different layers of a neural network over random data instances unrelated to the task being learned, via gradient ascent. The warped representations that result from this process are akin to the sensorial concoctions characteristic of real dreams, which are then interpreted by the brain in an attempt to make sense of their content. In the proposed technique, this is achieved via a soft-label optimization process whose goal is to minimize the impact of such new data over network parameters. Mixing of regular training data with few warped non-impactful instances is then theorized to reduce the likability of overfitting, with preliminary testing supporting this notion. Applicability of this algorithm is currently planned for several areas of AI research, to demonstrate its validity.

The original outline plan of this thesis envisioned exploring bio-inspired solutions for improved autonomy in deep learning, via emulation of emotional and plasticity processes of the brain. Overall it can be concluded that all goals set previously were achieved, resulting in meaningful contributions to the research community and betterment of AI towards greater emotional competence and adaptability.

10.2 Future Research

Works presented in this thesis are open-ended in that they may be further developed as standalone research, be combined for mutual improvement, or also provide insight and serve as starting points into broader topics within emotional competence and adaptability. Thus, the research directions proposed below address these three possibilities, ranging from concrete to more general goals for the future.

10.2.1 Emotion

As has been mentioned, a combination of works in emotion is often synergistic and yields improvement. Thus, it would be interesting to employ the principle presented in Chapter 5 regarding emotion fuzzification in an empathic experimental scenario as the one in Chapter 6. This can serve to understand whether empathy is only perceived from exact matches of emotion during an interaction or if its effects also occur when correlated states are demonstrated. For instance, a human-robot relationship may be strengthened by the artificial agent disliking or becoming angry at what made its user sad, considering how the same user could perceive such a reaction as the robot demonstrating concern for their well-being. An experimental procedure for this scenario could be similar to the one in Chapter 6. Yet instead, the ANN in charge of generating facial configurations for the artificial agent would include a fuzzification layer mid-way in its architecture so generation would consider emotional state correlations. Relatedly, training users would reward the agent when its facial expressions were somewhat related to the displayed emotional state, besides matching it perfectly. Finally, interaction with users could be evaluated in the same fashion to understand whether empathy would still be perceived, and results compared with those already published.

In terms of each work individually, and considering the growing adoption of dimensional approaches in emotional analysis, fuzzification of emotion dimensions could be explored to assess whether similar conclusions can be drawn compared to categorical techniques, both for recognition and expression. As for the work of Chapter 6, other emotional behaviors and social characteristics [281] should be explored depending on their usefulness for HRI. Projection, for example, could be explored to determine its detrimental effects on a relationship. Considering a more advanced robot system with emotion, it would be interesting to explore whether associating negative emotional states, caused by external factors, could also be detrimental to HRI despite users not being responsible for them. If so, this would mimic as well as provide valuable insight into human-human interaction.

Focusing on the bigger picture, research into emotionally competent AI must eventually adopt conventions regarding what is and/or what causes artificial emotion. This is paramount to allow the field of emotional expression to advance into naturally affective robotics, benefiting HRI, assistive services, and several other areas. One possibility is to draw inspiration from human development, in that emotion recognition and expression can result from a combination of continuous appraisals: internal state and peers' emotion, thus optimizing mimicry of states perceived as similar. Ergo, the following scheme could be used in future work, assuming an AI system whose internal variables constitute state r:

1. Scenario induces autonomic responses in agent coincidentally in state r;
2. Agent associates those responses and behavior with corresponding r;

- Recognition - Similar behavior detected in peers. Categorized as an expression of known state r;
- Expression - Re-occurrence of r due to external factors. Associated behavior used as a coping mechanism;

10.2.2 Adaptability

This concept is often addressed within the scope of specific abilities possessed by an artificial agent. As such, opportunities for research and development of AI adaptability are as numerous as AI characteristics and their applications. For instance, dopamine emulation could impact other aspects of learning besides convergence speed. This is because, while our results with neuromodulator emulation were underwhelming, other research has proved successful. Therefore, it would be fascinating to adapt the kind of signaling implemented in Chapter 7 to mediate engagement in a multi-task DL environment. Specifically, dopaminergic incentive salience could work to determine the order by which each stimulus is addressed. This would be achieved over recurrent contact with that environment, as greater dopamine release would be associated with preferred or more profitable tasks. Moreover, effects could be further expanded by integrating other neuromodulators, such as norepinephrine. Agents equipped with noradrenergic signaling could undergo arousal spikes, which could serve as drives to explore and seek out novel tasks within their environment. Several other chemicals (e.g. adrenaline, serotonin) could be considered and, similarly to Chapter 8, findings from these emulations could prove complementary to psychology and/or neuroscience studies.

Adaptive behavior is likewise a very broad topic. While Chapter 8 focuses on correlating emotion with exploration, the proposed framework retains the competence to study other behavioral traits and their relationship with internal drives. Mediating exploitation or engagement through variable emotion during cognitive operation could be a possibility, similar to how agents in that work learned to explore when seemingly more useful for their reward objective. On a separate note, Chapter 8 could also be extended to consider a more biologically plausible overlap of states. In this case, rather than considering emotions individually, several states could have a combined effect on general behavior. For instance, epistemic and achievement states could be employed in tandem, equitably, or via weighted contributions, as inputs to an actor module. This technique could entail either feature merging at a midstream level or an already multi-emotional input derived from a separate module.

In addition to the presented approaches, adaptability may also encompass dealing with unexpected issues during regular functioning. Lack of data, depletion of memory or computational resources, and overfitting all constitute examples of problems that are currently dealt with manually in AI methodology. Consequentially, it would be beneficial to research how these problems, or their biologically similar counterparts, are dealt with by the brain, to come up with possible solutions for artificial agents to achieve the same goal. Chapter 9 addressed this specifically, considering overfitting and presenting a way for neural networks to deal with the issue autonomously, inspired by a theory on how the brain addresses the same issue. While lacking thorough results, preliminary experimentation appeared promising, and so this work constitutes a solid road map for future research.

Bibliography

[1] J. M. Fellous, J. L. Armony, and J. E. LeDoux. "Emotional Circuits and Computational Neuroscience". In: *Neuroscience* 454.7200 (2002), pp. 1–8.

[2] C.M. Tyng et al. "The Influences of Emotion on Learning and Memory". In: *Frontiers in psychology* 8 (2017), p. 1454. DOI: https://doi.org/10.3389/fpsyg.2017.01454.

[3] Donna K. Housman. "The importance of emotional competence and self-regulation from birth: a case for the evidence-based emotional cognitive social early learning approach". In: *ICEP* 11.1 (Nov. 2017). DOI: 10.1186/s40723-017-0038-6. URL: https://doi.org/10.1186%2Fs40723-017-0038-6.

[4] M. Constandi. *Neuroplasticity.* Cambridge: The MIT Press: Essential Knowledge Series, Aug. 2016. ISBN: 978-0262529334.

[5] Penelope A. Lewis, Günther Knoblich, and Gina Poe. "How Memory Replay in Sleep Boosts Creative Problem-Solving". In: *Trends in Cognitive Sciences* 22.6 (June 2018), pp. 491–503. DOI: 10.1016/j.tics.2018.03.009. URL: https://doi.org/10.1016%2Fj.tics.2018.03.009.

[6] Marcin Pawel Jarzebski et al. "Ageing and population shrinking: implications for sustainability in the urban century". In: *npj Urban Sustain* 1.1 (May 2021). DOI: 10.1038/s42949-021-00023-z. URL: https://doi.org/10.1038%2Fs42949-021-00023-z.

[7] Eftychios G. Christoforou et al. "The Upcoming Role for Nursing and Assistive Robotics: Opportunities and Challenges Ahead". In: *Frontiers in Digital Health* 2 (2020). ISSN: 2673-253X. DOI: 10.3389/fdgth.2020.585656. URL: https://www.frontiersin.org/article/10.3389/fdgth.2020.585656.

[8] Ricardo Silva Peres et al. "Industrial Artificial Intelligence in Industry 4.0 - Systematic Review, Challenges and Outlook". In: *IEEE Access* 8 (2020), pp. 220121–220139. DOI: 10.1109/ACCESS.2020.3042874.

[9] Thomas B Sheridan. "A review of recent research in social robotics". In: *Current Opinion in Psychology* 36 (Dec. 2020), pp. 7–12. DOI: 10.1016/j.copsyc.2020.01.003. URL: https://doi.org/10.1016%2Fj.copsyc.2020.01.003.

[10] Ralph Adolphs. "How should neuroscience study emotions? by distinguishing emotion states, concepts, and experiences". In: *Social Cognitive and Affective Neuroscience* 12.1 (Oct. 2016), pp. 24–31. DOI: 10.1093/scan/nsw153. URL: https://doi.org/10.1093%2Fscan%2Fnsw153.

[11] Elmer Jacobs, Joost Broekens, and Catholijn M. Jonker. "Emergent Dynamics of Joy, Distress, Hope and Fear in Reinforcement Learning Agents". In: *Adaptive learning agents workshop at AAMAS2014*. 2014.

[12] D. Huggins-Daines et al. "Pocketsphinx: A Free, Real-Time Continuous Speech Recognition System for Hand-Held Devices". In: *2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings*. Vol. 1. 2006, pp. I–I. DOI: 10.1109/ICASSP.2006.1659988.

[13] *MATRIX Creator*. URL: matrix-io.github.io/matrix-documentation/matrix-creator/overview/.

[14] Gustavo Assunção, Nuno Gonçalves, and Paulo Menezes. "Bio-Inspired Modality Fusion for Active Speaker Detection". In: *Applied Sciences* 11.8 (2021). ISSN: 2076-3417. DOI: 10.3390/app11083397. URL: https://www.mdpi.com/2076-3417/11/8/3397.

[15] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain." In: *Psychological Review* 65.6 (1958), pp. 386–408. DOI: 10.1037/h0042519.

[16] Yoshua Bengio, Aaron Courville, and Pascal Vincent. "Representation Learning: A Review and New Perspectives". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.8 (2013), pp. 1798–1828. DOI: 10.1109/TPAMI.2013.50.

[17] Richard S. Sutton and Andrew G. Barto. *Reinforcement learning: An introduction*. MIT press, 2018.

[18] Christopher J. C. H. Watkins and Peter Dayan. "Q-learning". In: *Mach Learn* 8.3-4 (May 1992), pp. 279–292. DOI: 10.1007/bf00992698. URL: https://doi.org/10.1007%2Fbf00992698.

[19] Volodymyr Mnih et al. "Human-level control through deep reinforcement learning". In: *Nature* 518.7540 (Feb. 2015), pp. 529–533. DOI: 10.1038/nature14236. URL: https://doi.org/10.1038%2Fnature14236.

[20] Richard S Sutton et al. "Policy Gradient Methods for Reinforcement Learning with Function Approximation". In: *Advances in Neural Information Processing Systems*. Ed. by S. Solla, T. Leen, and K. Müller. Vol. 12. MIT Press, 1999.

[21] Thomas Degris, Martha White, and Richard S Sutton. "Linear Off-Policy Actor-Critic". In: *Proceedings of the International Conference on Machine Learning*. ICML'12. Edinburgh, Scotland, 2012.

[22] Timothy P. Lillicrap et al. "Continuous control with deep reinforcement learning". In: *4th International Conference on Learning Representations*. Ed. by Yoshua Bengio and Yann LeCun. 2016. URL: http://arxiv.org/abs/1509.02971.

[23] Gustavo Assuncao et al. "An Overview of Emotion in Artificial Intelligence". In: *IEEE Transactions on Artificial Intelligence* 3.6 (Dec. 2022), pp. 867–886. DOI: 10.1109/tai.2022.3159614. URL: https://doi.org/10.1109%2Ftai.2022.3159614.

[24] H. Lodish et al. *Molecular Cell Biology*. 9th. Section 21.1: Overview of Neuron Structure and Function. New York: New York: WH Freeman, 2021. ISBN: 9781319365028.

[25] Edmund T. Rolls. "Limbic systems for emotion and for memory, but no single limbic system". In: *Cortex* 62 (Jan. 2015), pp. 119–157. DOI: 10.1016/j.cortex.2013.12.005. URL: https://doi.org/10.1016%2Fj.cortex.2013.12.005.

[26] Joseph E. LeDoux. "Emotion Circuits in the Brain". In: *Annu. Rev. Neurosci.* 23.1 (Mar. 2000), pp. 155–184. DOI: 10.1146/annurev.neuro.23.1.155. URL: https://doi.org/10.1146%2Fannurev.neuro.23.1.155.

[27] Marcelo R. Roxo et al. "The Limbic System Conception and Its Historical Evolution". In: *The Scientific World JOURNAL* 11 (2011), pp. 2427–2440. DOI: 10.1100/2011/157150. URL: https://doi.org/10.1100%2F2011%2F157150.

[28] David M Schneider. "Reflections of action in sensory cortex". In: *Current Opinion in Neurobiology* 64 (Oct. 2020), pp. 53–59. DOI: 10.1016/j.conb.2020.02.004. URL: https://doi.org/10.1016%2Fj.conb.2020.02.004.

[29] Harald Stromfelt, Yue Zhang, and Bjorn W. Schuller. "Emotion-augmented machine learning: Overview of an emerging domain". In: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, Oct. 2017. DOI: 10.1109/acii.2017.8273617. URL: https://doi.org/10.1109%2Facii.2017.8273617.

[30] Elisabeth A. Murray. "The amygdala, reward and emotion". In: *Trends in Cognitive Sciences* 11.11 (Nov. 2007), pp. 489–497. DOI: 10.1016/j.tics.2007.08.013. URL: https://doi.org/10.1016%2Fj.tics.2007.08.013.

[31] Raphael Guex et al. "Temporal dynamics of amygdala response to emotion- and action-relevance". In: *Sci Rep* 10.1 (July 2020). DOI: 10.1038/s41598-020-67862-1. URL: https://doi.org/10.1038%2Fs41598-020-67862-1.

[32] Rudolf N. Cardinal et al. "Emotion and motivation: the role of the amygdala, ventral striatum, and prefrontal cortex". In: *Neuroscience & Biobehavioral Reviews* 26.3 (2002), pp. 321–352. ISSN: 0149-7634. DOI: https://doi.org/10.1016/S0149-7634(02)00007-6. URL: https://www.sciencedirect.com/science/article/pii/S0149763402000076.

[33] Patricia H. Janak and Kay M. Tye. "From circuits to behaviour in the amygdala". In: *Nature* 517.7534 (Jan. 2015), pp. 284–292. DOI: 10.1038/nature14188. URL: https://doi.org/10.1038%2Fnature14188.

[34] Elizabeth E. Steinberg et al. "Amygdala-Midbrain Connections Modulate Appetitive and Aversive Learning". In: *Neuron* 106.6 (June 2020), 1026–1043.e9. DOI: 10.1016/j.neuron.2020.03.016. URL: https://doi.org/10.1016%2Fj.neuron.2020.03.016.

[35] Michael Marxen et al. "Questioning the role of amygdala and insula in an attentional capture by emotional stimuli task". In: *Hum Brain Mapp* 42.5 (Nov. 2020), pp. 1257–1267. DOI: 10.1002/hbm.25290. URL: https://doi.org/10.1002%2Fhbm.25290.

[36] Katalin M. Gothard. "Multidimensional processing in the amygdala". In: *Nat Rev Neurosci* 21.10 (Aug. 2020), pp. 565–575. DOI: 10.1038/s41583-020-0350-y. URL: https://doi.org/10.1038%2Fs41583-020-0350-y.

[37] Richard Frackowiak. *Human brain function.* Amsterdam Boston: Elsevier Academic Press, 2004. ISBN: 9780080472959. DOI: 10.1016/b978-0-12-264841-0.x5000-8. URL: https://doi.org/10.1016%2Fb978-0-12-264841-0.x5000-8.

[38] Luiz Pessoa. "Emotion and cognition and the amygdala: From "what is it?" to "what's to be done?"" In: *Neuropsychologia* 48.12 (Oct. 2010), pp. 3416–3429. DOI: 10.1016/j.neuropsychologia.2010.06.038. URL: https://doi.org/10.1016%2Fj.neuropsychologia.2010.06.038.

[39] Edmund T. Rolls. *Emotion and Decision-making Explained.* Oxford University Press, Oct. 2013. DOI: 10.1093/acprof:oso/9780199659890.001.0001. URL: https://doi.org/10.1093%2Facprof%3Aoso%2F9780199659890.001.0001.

[40] Edmund T. Rolls. "The cingulate cortex and limbic systems for emotion, action, and memory". In: *Brain Struct Funct* 224.9 (Aug. 2019), pp. 3001–3018. DOI: 10.1007/s00429-019-01945-2. URL: https://doi.org/10.1007%2Fs00429-019-01945-2.

[41] Vani Pariyadath, Joshua L. Gowin, and Elliot A. Stein. "Resting state functional connectivity analysis for addiction medicine". In: *Progress in Brain Research.* Elsevier, 2016, pp. 155–173. DOI: 10.1016/bs.pbr.2015.07.015. URL: https://doi.org/10.1016%2Fbs.pbr.2015.07.015.

[42] Ja-Hyun Baik. "Stress and the dopaminergic reward system". In: *Experimental & Molecular Medicine* 52.12 (Dec. 2020), pp. 1879–1890. DOI: 10.1038/s12276-020-00532-4. URL: https://doi.org/10.1038%2Fs12276-020-00532-4.

[43] Mojtaba M. Asl, Abdol-Hossein Vahabie, and Alireza Valizadeh. "Dopaminergic Modulation of Synaptic Plasticity, Its Role in Neuropsychiatric Disorders, and Its Computational Modeling". In: *Basic Clin. Neurosci. J.* (Oct. 2018). DOI: 10.32598/bcn.9.10.125. URL: https://doi.org/10.32598%2Fbcn.9.10.125.

[44] Tiago V Maia and Michael J Frank. "From reinforcement learning models to psychiatric and neurological disorders". In: *Nat Neurosci* 14.2 (Jan. 2011), pp. 154–162. DOI: 10.1038/nn.2723. URL: https://doi.org/10.1038%2Fnn.2723.

[45] Kyle Dunovan and Timothy Verstynen. "Believer-Skeptic Meets Actor-Critic: Rethinking the Role of Basal Ganglia Pathways during Decision-Making and Reinforcement Learning". In: *Front. Neurosci.* 10 (Mar. 2016). DOI: 10.3389/fnins.2016.00106. URL: https://doi.org/10.3389%2Ffnins.2016.00106.

[46] Tai-Xiang Xu and Wei-Dong Yao. "D1 and D2 dopamine receptors in separate circuits cooperate to drive associative long-term potentiation in the prefrontal cortex". In: *Proc. Natl. Acad. Sci. U.S.A.* 107.37 (Aug. 2010), pp. 16366–16371. DOI: 10.1073/pnas.1004108107. URL: https://doi.org/10.1073%2Fpnas.1004108107.

[47] H. Trantham-Davidson. "Mechanisms Underlying Differential D1 versus D2 Dopamine Receptor Regulation of Inhibition in Prefrontal Cortex". In: *Journal of Neuroscience* 24.47 (Nov. 2004), pp. 10652–10659. DOI: 10.1523/jneurosci.3179-04.2004. URL: https://doi.org/10.1523%2Fjneurosci.3179-04.2004.

[48] V. Pawlak and J. N. D. Kerr. "Dopamine Receptor Activation Is Required for Corticostriatal Spike-Timing-Dependent Plasticity". In: *Journal of Neuroscience* 28.10 (Mar. 2008), pp. 2435–2446. DOI: 10.1523/jneurosci.4402-07.2008. URL: https://doi.org/10.1523%2Fjneurosci.4402-07.2008.

[49] R.G.M Morris. "D.O. Hebb: The Organization of Behavior, Wiley: New York, 1949". In: *Brain Research Bulletin* 50.5-6 (Nov. 1999), p. 437. DOI: 10.1016/s0361-9230(99)00182-3. URL: https://doi.org/10.1016/s0361-9230(99)00182-3.

[50] Wolfram Schultz. "Neuronal Reward and Decision Signals: From Theories to Data". In: *Physiological Reviews* 95.3 (July 2015), pp. 853–951. DOI: 10.1152/physrev.00023.2014. URL: https://doi.org/10.1152%2Fphysrev.00023.2014.

[51] Stephanie M. Groman. "The Neurobiology of Impulsive Decision-Making and Reinforcement Learning in Nonhuman Animals". In: *Recent Advances in Research on Impulsivity and Impulsive Behaviors*. Springer International Publishing, 2020, pp. 23–52. DOI: 10.1007/7854_2020_127. URL: https://doi.org/10.1007%2F7854_2020_127.

[52] Abbas Khani and Gregor Rainer. "Neural and neurochemical basis of reinforcement guided decision making". In: *Journal of Neurophysiology* 116.2 (Aug. 2016), pp. 724–741. DOI: 10.1152/jn.01113.2015. URL: https://doi.org/10.1152%2Fjn.01113.2015.

[53] Fushun Wang et al. "Editorial: Neurotransmitters and Emotions". In: *Front. Psychol.* 11 (Jan. 2020). DOI: 10.3389/fpsyg.2020.00021. URL: https://doi.org/10.3389%2Ffpsyg.2020.00021.

[54] Christopher D Fiorillo, Philippe N Tobler, and Wolfram Schultz. "Evidence that the delay-period activity of dopamine neurons corresponds to reward uncertainty rather than backpropagating TD errors". In: *Behav Brain Funct* 1.1 (June 2005). DOI: 10.1186/1744-9081-1-7. URL: https://doi.org/10.1186%2F1744-9081-1-7.

[55] G Aston-Jones et al. "Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task". In: *J. Neurosci.* 14.7 (July 1994), pp. 4467–4480. DOI: 10.1523/jneurosci.14-07-04467.1994. URL: https://doi.org/10.1523%2Fjneurosci.14-07-04467.1994.

[56] Trond Myhrer. "Neurotransmitter systems involved in learning and memory in the rat: a meta-analysis based on studies of four behavioral tasks". In: *Brain Research Reviews* 41.2 (2003), pp. 268–287. ISSN: 0165-0173. DOI: https://doi.org/10.1016/S0165-0173(02)00268-0. URL: https://www.sciencedirect.com/science/article/pii/S0165017302002680.

[57] Lisa Feldman Barrett. "2 - Navigating the Science of Emotion". In: *Emotion Measurement*. Ed. by Herbert L. Meiselman. Woodhead Publishing, 2016, pp. 31–63. ISBN: 978-0-08-100508-8. DOI: https://doi.org/10.1016/B978-0-08-100508-8.00002-3. URL: https://www.sciencedirect.com/science/article/pii/B9780081005088000023.

[58] Agnes Moors et al. "Appraisal Theories of Emotion: State of the Art and Future Development". In: *Emotion Review* 5.2 (Mar. 2013), pp. 119–124. DOI: 10.1177/1754073912468165. URL: https://doi.org/10.1177%2F1754073912468165.

[59] Dacher Keltner and James J. Gross. "Functional Accounts of Emotions". In: *Cognition and Emotion* 13.5 (Sept. 1999), pp. 467–480. DOI: 10.1080/026999399379140. URL: https://doi.org/10.1080%2F026999399379140.

[60] Richard S. Lazarus. "Emotion and Adaptation". In: vol. 21. Apr. 1991, p. 557. ISBN: 9780195092660.

[61] Randy J. Larsen and Zvjezdana Prizmic-Larsen. "Measuring Emotions: Implications of a Multimethod Perspective." In: *Handbook of multimethod measurement in psychology*. American Psychological Association, 2006, pp. 337–351. DOI: 10.1037/11383-023. URL: https://doi.org/10.1037%2F11383-023.

[62] R. W. Levenson. "The Nature of Emotion: Fundamental Questions". In: Oxford University Press, 1994. Chap. Human Emotions: A Functional View, pp. 123–126. ISBN: 9780195089448.

[63] Antonio Damasio and Gil B. Carvalho. "The nature of feelings: evolutionary and neurobiological origins". In: *Nat Rev Neurosci* 14.2 (Jan. 2013), pp. 143–152. DOI: 10.1038/nrn3403. URL: https://doi.org/10.1038%2Fnrn3403.

[64] James J. Gross. "Emotion Regulation: Past, Present, Future". In: *Cognition and Emotion* 13.5 (Sept. 1999), pp. 551–573. DOI: 10.1080/026999399379186. URL: https://doi.org/10.1080%2F026999399379186.

[65] Laith Al-Shawaf et al. "Human Emotions: An Evolutionary Psychological Perspective". In: *Emotion Review* 8.2 (Feb. 2015), pp. 173–186. DOI: 10.1177/1754073914565518.

[66] C. Smith and R. Lazarus. "Handbook of personality: theory and research". In: 2nd. 1990. Chap. Emotion and Adaptation, pp. 609–637. ISBN: 1-57230-483-9.

[67] James J. Gross. "Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology." In: *Journal of Personality and Social Psychology* 74.1 (1998), pp. 224–237. DOI: 10.1037/0022-3514.74.1.224. URL: https://doi.org/10.1037/0022-3514.74.1.224.

[68] Stephen W Porges. *The polyvagal theory: Neurophysiological foundations of emotions, attachment, communication, and self-regulation.* Norton Series on Interpersonal Neurobiology. New York, NY: WW Norton, Apr. 2011.

[69] Joseph LeDoux. "The emotional brain, fear, and the amygdala". In: *Cellular and molecular neurobiology* 23.4 (2003), pp. 727–738.

[70] Stephanie Burnett et al. "Development during Adolescence of the Neural Processing of Social Emotion". In: *Journal of Cognitive Neuroscience* 21.9 (Sept. 2009), pp. 1736–1750. ISSN: 0898-929X. DOI: 10.1162/jocn.2009.21121.

[71] Mary M Herrald and Joe Tomaka. "Patterns of emotion-specific appraisal, coping, and cardiovascular reactivity during an ongoing emotional episode." In: *Journal of personality and social psychology* 83.2 (2002), p. 434.

[72] Jeremy P. Jamieson, Matthew K. Nock, and Wendy Berry Mendes. "Mind over matter: Reappraising arousal improves cardiovascular and cognitive responses to stress." In: *Journal of Experimental Psychology: General* 141.3 (2012), pp. 417–422. DOI: 10.1037/a0025719. URL: https://doi.org/10.1037%2Fa0025719.

[73] Robert Plutchik. "The Nature of Emotions: Human Emotions Have Deep Evolutionary Roots, a Fact That May Explain Their Complexity and Provide Tools for Clinical Practice". In: 4. JSTOR, 2001, 344–350. URL: http://www.jstor.org/stable/27857503.

[74] Paul Ekman. "An argument for basic emotions". In: *Cognition and Emotion* 6.3-4 (1992), pp. 169–200. DOI: 10.1080/02699939208411068.

[75] Peter J Lang. "The emotion probe: Studies of motivation and attention." In: *American psychologist* 50.5 (1995), p. 372.

[76] Albert Mehrabian. "Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression". In: *Journal of psychopathology and behavioral assessment* 19.4 (1997), pp. 331–357.

[77] Marianne Chevrier et al. "Exploring the antecedents and consequences of epistemic emotions". In: *Learning and Instruction* 63 (Oct. 2019), p. 101209. DOI: 10.1016/j.learninstruc.2019.05.006. URL: https://doi.org/10.1016%2Fj.learninstruc.2019.05.006.

[78] S. A. Fitneva and M. Slinger. "Looking for a second opinion: Epistemic emotions and the exploration of information sources". In: *Proceedings of the Annual Meeting of the Cognitive Science Society* (2022). URL: https://escholarship.org/uc/item/60c7n8k8.

[79] Elisabeth Vogl et al. "Surprise, Curiosity, and Confusion Promote Knowledge Exploration: Evidence for Robust Effects of Epistemic Emotions". In: *Frontiers in Psychology* 10 (2019). ISSN: 1664-1078. DOI: `10.3389/fpsyg.2019.02474`. URL: `https://www.frontiersin.org/article/10.3389/fpsyg.2019.02474`.

[80] Elisabeth Vogl et al. "Surprised–curious–confused: Epistemic emotions and knowledge exploration". In: *Emotion* 20.4 (June 2020), pp. 625–641. DOI: `10.1037/emo0000578`. URL: `https://doi.org/10.1037%2Femo0000578`.

[81] Sidney D'Mello and Art Graesser. "Confusion and its dynamics during device comprehension with breakdown scenarios". en. In: *Acta Psychol. (Amst.)* 151 (Sept. 2014), pp. 106–116.

[82] K Ann Renninger and Suzanne E Hidi. *The power of interest for motivation and engagement.* and Suzanne Hidi. Description: New York, NY : Routledge, 2016. —: Routledge, Nov. 2015.

[83] Lisa A Williams and David DeSteno. "Pride and perseverance: the motivational role of pride". en. In: *J. Pers. Soc. Psychol.* 94.6 (June 2008), pp. 1007–1017.

[84] Elaine Hatfield, John T. Cacioppo, and Richard L. Rapson. "Emotional Contagion". In: *Current Directions in Psychological Science* 2.3 (1993), pp. 96–99. ISSN: 09637214. URL: `http://www.jstor.org/stable/20182211` (visited on 04/27/2022).

[85] Amit Goldenberg and James J. Gross. "Digital Emotion Contagion". In: *Trends in Cognitive Sciences* 24.4 (2020), pp. 316–328. ISSN: 1364-6613. DOI: `https://doi.org/10.1016/j.tics.2020.01.009`. URL: `https://www.sciencedirect.com/science/article/pii/S1364661320300279`.

[86] Carolina Herrando and Efthymios Constantinides. "Emotional Contagion: A Brief Overview and Future Directions". In: *Frontiers in Psychology* 12 (2021). ISSN: 1664-1078. DOI: `10.3389/fpsyg.2021.712606`.

[87] Sigal G. Barsade, Constantinos G.V. Coutifaris, and Julianna Pillemer. "Emotional contagion in organizational life". In: *Research in Organizational Behavior* 38 (2018), pp. 137–151. ISSN: 0191-3085. DOI: `https://doi.org/10.1016/j.riob.2018.11.005`. URL: `https://www.sciencedirect.com/science/article/pii/S0191308518300108`.

[88] Martin L. Hoffman. *Empathy and Moral Development: Implications for Caring and Justice.* Cambridge University Press, 2000.

[89] Ezra Stotland. "Exploratory Investigations of Empathy11The preparation of this article and all of the initially reported studies were supported by a grant from the National Science Foundation." In: ed. by Leonard Berkowitz. Vol. 4. Advances in Experimental Social Psychology. Academic Press, 1969, pp. 271–314. DOI: `https://doi.org/10.1016/S0065-2601(08)60080-5`. URL: `https://www.sciencedirect.com/science/article/pii/S0065260108600805`.

[90] Elahe Bagheri et al. "A Reinforcement Learning Based Cognitive Empathy Framework for Social Robots". In: *Int J of Soc Robotics* 13.5 (Sept. 2020), pp. 1079–1093. DOI: 10.1007/s12369-020-00683-4. URL: https://doi.org/10.1007%2Fs12369-020-00683-4.

[91] Stela H. Seo et al. "Poor Thing! Would You Feel Sorry for a Simulated Robot? A comparison of empathy toward a physical and a simulated robot". In: *2015 10th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 2015, pp. 125–132.

[92] Carolyn Saarni. *The development of emotional competence*. Guilford press, 1999.

[93] Velvetina Lim, Maki Rooksby, and Emily S. Cross. "Social Robots on a Global Stage: Establishing a Role for Culture During Human–Robot Interaction". In: *International Journal of Social Robotics* 13.6 (Nov. 2020), pp. 1307–1333. DOI: 10.1007/s12369-020-00710-4. URL: https://doi.org/10.1007%2Fs12369-020-00710-4.

[94] Maike Paetzel-Prüsmann, Giulia Perugia, and Ginevra Castellano. "The Influence of robot personality on the development of uncanny feelings". In: *Computers in Human Behavior* 120 (July 2021), p. 106756. DOI: 10.1016/j.chb.2021.106756. URL: https://doi.org/10.1016%2Fj.chb.2021.106756.

[95] Dimitrios Koutentakis, Alexander Pilozzi, and Xudong Huang. "Designing Socially Assistive Robots for Alzheimer's Disease and Related Dementia Patients and Their Caregivers: Where We Are and Where We Are Headed". In: *Healthcare* 8.2 (Mar. 2020), p. 73. DOI: 10.3390/healthcare8020073. URL: https://doi.org/10.3390%2Fhealthcare8020073.

[96] Ester Martinez-Martin, Felix Escalona, and Miguel Cazorla. "Socially Assistive Robots for Older Adults and People with Autism: An Overview". In: *Electronics* 9.2 (Feb. 2020), p. 367. DOI: 10.3390/electronics9020367. URL: https://doi.org/10.3390%2Felectronics9020367.

[97] Ruud Hortensius, Felix Hekele, and Emily S. Cross. "The Perception of Emotion in Artificial Agents". In: *IEEE Transactions on Cognitive and Developmental Systems* 10.4 (Dec. 2018), pp. 852–864. DOI: 10.1109/tcds.2018.2826921. URL: https://doi.org/10.1109%2Ftcds.2018.2826921.

[98] Naima Otberdout et al. "Automatic Analysis of Facial Expressions Based on Deep Covariance Trajectories". In: *IEEE Transactions on Neural Networks and Learning Systems* 31.10 (Oct. 2020), pp. 3892–3905. DOI: 10.1109/tnnls.2019.2947244. URL: https://doi.org/10.1109%2Ftnnls.2019.2947244.

[99] Bilal Taha et al. "Learned 3D Shape Representations Using Fused Geometrically Augmented Images: Application to Facial Expression and Action Unit Detection". In: *IEEE Transactions on Circuits and Systems for Video Technology* 30.9 (Sept. 2020), pp. 2900–2916. DOI: 10.1109/tcsvt.2020.2984241. URL: https://doi.org/10.1109%2Ftcsvt.2020.2984241.

[100] Anis Kacem et al. "A Novel Geometric Framework on Gram Matrix Trajectories for Human Behavior Understanding". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42.1 (Jan. 2020), pp. 1–14. DOI: 10.1109/tpami.2018.2872564. URL: https://doi.org/10.1109%2Ftpami.2018.2872564.

[101] Weijian Zhang and Peng Song. "Transfer Sparse Discriminant Subspace Learning for Cross-Corpus Speech Emotion Recognition". In: *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 28 (2020), pp. 307–318. DOI: 10.1109/taslp.2019.2955252. URL: https://doi.org/10.1109%2Ftaslp.2019.2955252.

[102] Zhaocheng Huang and Julien Epps. "An Investigation of Partition-Based and Phonetically-Aware Acoustic Features for Continuous Emotion Prediction from Speech". In: *IEEE Transactions on Affective Computing* 11.4 (Oct. 2020), pp. 653–668. DOI: 10.1109/taffc.2018.2821135. URL: https://doi.org/10.1109%2Ftaffc.2018.2821135.

[103] Tengfei Song et al. "EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks". In: *IEEE Transactions on Affective Computing* 11.3 (July 2020), pp. 532–541. DOI: 10.1109/taffc.2018.2817622. URL: https://doi.org/10.1109%2Ftaffc.2018.2817622.

[104] J.A. Dominguez-Jimenez et al. "A machine learning model for emotion recognition from physiological signals". In: *Biomedical Signal Processing and Control* 55 (Jan. 2020), p. 101646. DOI: 10.1016/j.bspc.2019.101646. URL: https://doi.org/10.1016%2Fj.bspc.2019.101646.

[105] Niko Colneric and Janez Demsar. "Emotion Recognition on Twitter: Comparative Study and Training a Unison Model". In: *IEEE Transactions on Affective Computing* 11.3 (July 2020), pp. 433–446. DOI: 10.1109/taffc.2018.2807817. URL: https://doi.org/10.1109%2Ftaffc.2018.2807817.

[106] Luke Hickton, Matthew Lewis, and Lola Cañamero. "Expression of Grounded Affect: How Much Emotion Can Arousal Convey?" In: *Towards Autonomous Robotic Systems*. Springer International Publishing, 2020, pp. 234–248. DOI: 10.1007/978-3-030-63486-5_26. URL: https://doi.org/10.1007%2F978-3-030-63486-5_26.

[107] Xiqian Zheng et al. "What Kinds of Robot's Touch Will Match Expressed Emotions?" In: *IEEE Robotics and Automation Letters* 5.1 (Jan. 2020), pp. 127–134. DOI: 10.1109/lra.2019.2947010. URL: https://doi.org/10.1109%2Flra.2019.2947010.

[108] Mathias Sunardi and Marek Perkowski. "Behavior Expressions for Social and Entertainment Robots". In: *2020 IEEE 50th International Symposium on Multiple-Valued Logic (ISMVL)*. IEEE, Nov. 2020. DOI: 10.1109/ismvl49045.2020.00058. URL: https://doi.org/10.1109/ismvl49045.2020.00058.

[109] Yusuke Nishimura, Yutaka Nakamura, and Hiroshi Ishiguro. "Human interaction behavior modeling using Generative Adversarial Networks". In: *Neural Networks* 132 (Dec. 2020), pp. 521–531. DOI: 10.1016/j.neunet.2020.09.019. URL: https://doi.org/10.1016%2Fj.neunet.2020.09.019.

[110] Nguyen Tan Viet Tuyen, Armagan Elibol, and Nak Young Chong. “Learning Bodily Expression of Emotion for Social Robots Through Human Interaction”. In: *IEEE Transactions on Cognitive and Developmental Systems* 13.1 (Mar. 2021), pp. 16–30. DOI: `10.1109/tcds.2020.3005907`. URL: `https://doi.org/10.1109/tcds.2020.3005907`.

[111] Lu Chen et al. “Automatic Chinese Font Generation System Reflecting Emotions Based on Generative Adversarial Network”. In: *Applied Sciences* 10.17 (Aug. 2020), p. 5976. DOI: `10.3390/app10175976`. URL: `https://doi.org/10.3390%2Fapp10175976`.

[112] Mauajama Firdaus et al. “EmoSen: Generating Sentiment and Emotion Controlled Responses in a Multimodal Dialogue System”. In: *IEEE Transactions on Affective Computing* 13.3 (July 2022), pp. 1555–1566. DOI: `10.1109/taffc.2020.3015491`. URL: `https://doi.org/10.1109%2Ftaffc.2020.3015491`.

[113] Klaus Weber et al. “How to Shape the Humor of a Robot - Social Behavior Adaptation Based on Reinforcement Learning”. In: *Proceedings of the 20th ACM International Conference on Multimodal Interaction*. ACM, Oct. 2018. DOI: `10.1145/3242969.3242976`. URL: `https://doi.org/10.1145%2F3242969.3242976`.

[114] Rui Liu, Berrak Sisman, and Haizhou Li. “Reinforcement Learning for Emotional Text-to-Speech Synthesis with Improved Emotion Discriminability”. In: 2104.01408 (2021).

[115] Neziha Akalin and Amy Loutfi. “Reinforcement Learning Approaches in Social Robotics”. In: *Sensors* 21.4 (2021). DOI: `10.3390/s21041292`.

[116] Ha-Duong Bui and Nak Young Chong. “Autonomous Speech Volume Control for Social Robots in a Noisy Environment Using Deep Reinforcement Learning”. In: *2019 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. 2019, pp. 1263–1268. DOI: `10.1109/ROBIO49542.2019.8961810`.

[117] A Castro-Gonzalez, M Malfaz, and M A Salichs. “An Autonomous Social Robot in Fear”. In: *IEEE Trans. Auton. Ment. Dev.* 5.2 (June 2013), pp. 135–151. DOI: `10.1109/TAMD.2012.2234120`.

[118] Henry Williams et al. “Emotion inspired adaptive robotic path planning”. In: *2015 IEEE Congress on Evolutionary Computation (CEC)*. IEEE, May 2015. DOI: `10.1109/cec.2015.7257263`. URL: `https://doi.org/10.1109%2Fcec.2015.7257263`.

[119] Panos Achlioptas et al. “ArtEmis: Affective Language for Visual Art”. In: *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021, pp. 11564–11574. DOI: `10.1109/CVPR46437.2021.01140`.

[120] Angelica Lim and Hiroshi G. Okuno. “The MEI Robot: Towards Using Motherese to Develop Multimodal Emotional Intelligence”. In: *IEEE Transactions on Autonomous Mental Development* 6.2 (June 2014), pp. 126–138. DOI: `10.1109/tamd.2014.2317513`. URL: `https://doi.org/10.1109%2Ftamd.2014.2317513`.

[121] Seth D. Pollak, Linda A. Camras, and Pamela M. Cole. "Progress in understanding the emergence of human emotion." In: *Developmental Psychology* 55.9 (Sept. 2019), pp. 1801–1811. DOI: 10.1037/dev0000789. URL: https://doi.org/10.1037%2Fdev0000789.

[122] Thomas Riccio. "Sophia Robot". In: *TDR: The Drama Review* 65.3 (Sept. 2021), pp. 42–77. DOI: 10.1017/s1054204321000319. URL: https://doi.org/10.1017%2Fs1054204321000319.

[123] Shuichi Nishio, Hiroshi Ishiguro, and Norihiro Hagit. "Geminoid: Teleoperated Android of an Existing Person". In: *Humanoid Robots: New Developments*. I-Tech Education and Publishing, June 2007. DOI: 10.5772/4876. URL: https://doi.org/10.5772%2F4876.

[124] Luefeng Chen et al. "Emotional Human-Robot Interaction Systems". In: *Emotion Recognition and Understanding for Emotional Human-Robot Interaction Systems*. Springer International Publishing, Nov. 2020, pp. 215–222. DOI: 10.1007/978-3-030-61577-2_12. URL: https://doi.org/10.1007%2F978-3-030-61577-2_12.

[125] Kwadwo Opong-Mensah. "Simulation of Human and Artificial Emotion (SHArE)". In: 2011.02151 (2020).

[126] John E. Steephen et al. "HED-ID: An Affective Adaptation Model Explaining the Intensity-Duration Relationship of Emotion". In: *IEEE Transactions on Affective Computing* 11.4 (Oct. 2020), pp. 736–750. DOI: 10.1109/taffc.2018.2848656. URL: https://doi.org/10.1109%2Ftaffc.2018.2848656.

[127] Suman Ojha, Jonathan Vitale, and Mary-Anne Williams. "Computational emotion models: A thematic review". en. In: *Int. J. Soc. Robot.* 13.6 (Sept. 2021), pp. 1253–1279. DOI: 10.1007/s12369-020-00713-1.

[128] Enrique Osuna, Luis-Felipe Rodríguez, and J Octavio Gutierrez-Garcia. "Toward integrating cognitive components with computational models of emotion using software design patterns". en. In: *Cogn. Syst. Res.* 65 (Jan. 2021), pp. 138–150. DOI: 10.1016/j.cogsys.2020.10.004.

[129] Enrique Osuna et al. "Development of computational models of emotions: A software engineering perspective". en. In: *Cogn. Syst. Res.* 60 (May 2020), pp. 1–19. DOI: 10.1016/j.cogsys.2019.11.001.

[130] Pablo Brinol, Richard E Petty, and Derek D Rucker. "The role of metacognitive processes in emotional intelligence". en. In: *Psicothema* 18 Suppl (2006), pp. 26–33.

[131] K. Richard Ridderinkhof. "Emotion in Action: A Predictive Processing Perspective and Theoretical Synthesis". In: *Emotion Review* 9.4 (Aug. 2017), pp. 319–325. DOI: 10.1177/1754073916661765. URL: https://doi.org/10.1177%2F1754073916661765.

[132] V. Mäki-Marttunen, T. Hagen, and T. Espeseth. "Proactive and reactive modes of cognitive control can operate independently and simultaneously". In: *Acta Psychologica* 199 (Aug. 2019), p. 102891. DOI: 10.1016/j.actpsy.2019.102891. URL: https://doi.org/10.1016%2Fj.actpsy.2019.102891.

[133] Gustavo Assunção and Paulo Menezes. "Intermediary Fuzzification in Speech Emotion Recognition". In: *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, July 2020, pp. 1–6. DOI: 10.1109/FUZZ48607.2020.9177699.

[134] Anna Esposito et al. "Seniors' Appreciation of Humanoid Robots". In: *Neural Approaches to Dynamics of Signal Exchanges*. Ed. by Anna Esposito et al. Singapore: Springer Singapore, 2020, pp. 331–345. ISBN: 978-981-13-8950-4. DOI: 10.1007/978-981-13-8950-4_30.

[135] Berardina De Carolis, Stefano Ferilli, and Giuseppe Palestra. "Simulating empathic behavior in a social assistive robot". In: *Multimed Tools Appl* 76.4 (Sept. 2016), pp. 5073–5094. DOI: 10.1007/s11042-016-3797-0. URL: https://doi.org/10.1007%2Fs11042-016-3797-0.

[136] Peng Song and Wenming Zheng. "Feature selection based transfer subspace learning for speech emotion recognition". In: *IEEE Trans. Affect. Comput.* 11.3 (July 2020), pp. 373–382.

[137] Huan Zhao, Yufeng Xiao, and Zixing Zhang. "Robust Semisupervised Generative Adversarial Networks for Speech Emotion Recognition via Distribution Smoothness". In: *IEEE Access* 8 (2020), pp. 106889–106900. DOI: 10.1109/ACCESS.2020.3000751.

[138] Mustaqeem, Muhammad Sajjad, and Soonil Kwon. "Clustering-based speech emotion recognition by incorporating learned features and deep BiLSTM". In: *IEEE Access* 8 (2020), pp. 79861–79875.

[139] Leila Kerkeni et al. "Automatic speech emotion recognition using machine learning". In: *Social Media and Machine Learning*. IntechOpen, Feb. 2020.

[140] Siddique Latif et al. "Transfer learning for improving speech emotion classification accuracy". In: *Interspeech 2018*. ISCA: ISCA, Sept. 2018.

[141] Geoffrey Hinton. "Deep belief networks". In: *Scholarpedia* 4.5 (2009), p. 5947. DOI: 10.4249/scholarpedia.5947. URL: https://doi.org/10.4249/scholarpedia.5947.

[142] Randy J Larsen and Edward Diener. "Promises and problems with the circumplex model of emotion." In: (1992).

[143] Xiang Pan. "Research on the Emotion Recognition Based on the Fuzzy Neural Network in the Intelligence Education System". In: *2011 Second International Conference on Digital Manufacturing & Automation*. IEEE, Aug. 2011. DOI: 10.1109/icdma.2011.255. URL: https://doi.org/10.1109%2Ficdma.2011.255.

[144] Wenhui Shi and Mingyan Jiang. "Fuzzy Wavelet Network with Feature Fusion and LM Algorithm for Facial Emotion Recognition". In: *2018 IEEE International Conference of Safety Produce Informatization (IICSPI)*. 2018, pp. 582–586. DOI: 10.1109/IICSPI.2018.8690353.

[145] E. Lotfi, A. Khosravi, and S. Nahavandi. "Facial emotion recognition using emotional neural network and hybrid of fuzzy c-means and genetic algorithm". In: *2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, July 2017. DOI: 10.1109/fuzz-ieee.2017.8015591. URL: https://doi.org/10.1109%2Ffuzz-ieee.2017.8015591.

[146] V. V. Borisov and K. P. Korshunova. "Multiclass Classification Based on the Convolutional Fuzzy Neural Networks". In: *Communications in Computer and Information Science*. Springer International Publishing, 2019, pp. 226–233. DOI: 10.1007/978-3-030-30763-9_19. URL: https://doi.org/10.1007%2F978-3-030-30763-9_19.

[147] Arsha Nagrani, Joon Son Chung, and Andrew Zisserman. "VoxCeleb: A Large-Scale Speaker Identification Dataset". In: *Proc. Interspeech 2017*. 2017, pp. 2616–2620. DOI: 10.21437/Interspeech.2017-950.

[148] Gustavo Assunção, Paulo Menezes, and Fernando Perdigão. "Speaker Awareness for Speech Emotion Recognition". In: *Int. J. Onl. Eng.* 16.04 (Apr. 2020), p. 15. DOI: 10.3991/ijoe.v16i04.11870. URL: https://doi.org/10.3991%2Fijoe.v16i04.11870.

[149] Joon Son Chung, Arsha Nagrani, and Andrew Zisserman. "VoxCeleb2: Deep Speaker Recognition". In: *Proc. Interspeech 2018*. 2018, pp. 1086–1090. DOI: 10.21437/Interspeech.2018-1929.

[150] James C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Springer US, 1981. DOI: 10.1007/978-1-4757-0450-1. URL: https://doi.org/10.1007%2F978-1-4757-0450-1.

[151] Felix Burkhardt et al. "A database of German emotional speech". In: *Interspeech 2005*. ISCA, Sept. 2005. DOI: 10.21437/interspeech.2005-446. URL: https://doi.org/10.21437%2Finterspeech.2005-446.

[152] Giovanni Costantini et al. "EMOVO Corpus: an Italian Emotional Speech Database". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Reykjavik, Iceland: European Language Resources Association (ELRA), May 2014, pp. 3501–3504. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/591_Paper.pdf.

[153] Sanaul Haq, Philip J. B. Jackson, and James D. Edge. "Audio-visual feature selection and reduction for emotion classification". In: *Proc. Int'l Conf. on Auditory-Visual Speech Processing*. 2008, pp. 185–190.

[154] ELRA. *Emotional speech synthesis database S0329*. catalogue.elra.info/en-us/repository/browse/ELRA-S0329/. 2012.

[155] Maxim Sidorov, Stefan Ultes, and Alexander Schmitt. "Emotions are a personal thing: Towards speaker-adaptive emotion recognition". In: *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2014, pp. 4803–4807. DOI: 10.1109/ICASSP.2014.6854514.

[156] Alessandra Sorrentino et al. "Modeling affective empathy by teaching emotion expressions to a social robot". In: *Workshop: Social Robots for Personalized, Continuous and Adaptive Assistance (ALTRUIST 2021), 13th International Conference on Social Robotics (ICSR)*. 2021. URL: http://altruist21.istc.cnr.it/.

[157] Alessandra Sorrentino et al. "A Reinforcement Learning Framework to Foster Affective Empathy in Social Robots". In: *Social Robotics*. Springer Nature Switzerland, 2022, pp. 522–533. DOI: 10.1007/978-3-031-24667-8_46. URL: https://doi.org/10.1007%2F978-3-031-24667-8_46.

[158] Berat A. Erol et al. "Toward Artificial Emotional Intelligence for Cooperative Social Human-Machine Interaction". In: *IEEE Transactions on Computational Social Systems* 7.1 (Feb. 2020), pp. 234–246. DOI: 10.1109/tcss.2019.2922593. URL: https://doi.org/10.1109%2Ftcss.2019.2922593.

[159] Minoru Asada. "Artificial Pain May Induce Empathy, Morality, and Ethics in the Conscious Mind of Robots". In: *Philosophies* 4.3 (July 2019), p. 38. DOI: 10.3390/philosophies4030038. URL: https://doi.org/10.3390%2Fphilosophies4030038.

[160] Mahni Shayganfar et al. ""It Was Not Your Fault" – Emotional Awareness Improves Collaborative Robots". In: *2019 IEEE International Conference on Humanized Computing and Communication (HCC)*. 2019, pp. 7–15. DOI: 10.1109/HCC46620.2019.00010.

[161] Iolanda Leite et al. "The influence of empathy in human–robot relations". In: *International Journal of Human-Computer Studies* 71.3 (2013), pp. 250–260. ISSN: 1071-5819. DOI: https://doi.org/10.1016/j.ijhcs.2012.09.005. URL: https://www.sciencedirect.com/science/article/pii/S1071581912001681.

[162] Andreea Niculescu et al. "Making Social Robots More Attractive: The Effects of Voice Pitch, Humor and Empathy". In: *Int J of Soc Robotics* 5.2 (Jan. 2013), pp. 171–191. DOI: 10.1007/s12369-012-0171-x. URL: https://doi.org/10.1007%2Fs12369-012-0171-x.

[163] Frans B. M. de Waal and Stephanie D. Preston. "Mammalian empathy: behavioural manifestations and neural basis". In: *Nat Rev Neurosci* 18.8 (June 2017), pp. 498–509. DOI: 10.1038/nrn.2017.72. URL: https://doi.org/10.1038%2Fnrn.2017.72.

[164] Ana Paiva et al. "Empathy in virtual agents and robots: A survey". In: *ACM Transactions on Interactive Intelligent Systems (TiiS)* 7.3 (2017), pp. 1–40.

[165] Barbara Gonsior et al. "Improving aspects of empathy and subjective performance for HRI through mirroring facial expressions". In: *2011 RO-MAN*. IEEE, 2011, pp. 350–356. DOI: 10.1109/ROMAN.2011.6005294.

[166] Laurel D. Riek, Philip C. Paul, and Peter Robinson. "When my robot smiles at me: Enabling human-robot rapport via real-time head gesture mimicry". In: *J Multimodal User Interfaces* 3.1-2 (Nov. 2009), pp. 99–108. DOI: 10.1007/s12193-009-0028-2. URL: https://doi.org/10.1007%2Fs12193-009-0028-2.

[167] Hifza Javed and Chung Hyuk Park. "Interactions With an Empathetic Agent: Regulating Emotions and Improving Engagement in Autism". In: *IEEE Robotics and Automation Magazine* 26.2 (June 2019), pp. 40–48. DOI: 10.1109/mra.2019.2904638. URL: https://doi.org/10.1109%2Fmra.2019.2904638.

[168] Nikhil Churamani et al. "Teaching emotion expressions to a human companion robot using deep neural architectures". In: *2017 International Joint Conference on Neural Networks (IJCNN)*. 2017, pp. 627–634. DOI: 10.1109/IJCNN.2017.7965911.

[169] Nikhil Churamani et al. "Learning Empathy-Driven Emotion Expressions using Affective Modulations". In: *2018 International Joint Conference on Neural Networks (IJCNN)*. 2018, pp. 1–8. DOI: 10.1109/IJCNN.2018.8489158.

[170] David Portugal et al. "A Study on the Deployment of a Service Robot in an Elderly Care Center". In: *Int J of Soc Robotics* 11.2 (Nov. 2018), pp. 317–341. DOI: 10.1007/s12369-018-0492-5. URL: https://doi.org/10.1007%2Fs12369-018-0492-5.

[171] Paul Ekman and Wallace V. Friesen. "Constants across cultures in the face and emotion." In: *Journal of Personality and Social Psychology* 17.2 (1971), pp. 124–129. DOI: 10.1037/h0030377. URL: https://doi.org/10.1037%2Fh0030377.

[172] Alessandra Sorrentino. "From humans to robots: leveraging social intelligence to improve HRI". PhD thesis. Pisa, Italy: Scuola Superiore Sant'Anna, 2022.

[173] P. Lang, M. Bradley, and B. Cuthbert. "International affective picture system (IAPS): Affective ratings of pictures and instruction manual". In: *Technical Report A-8* (2008). URL: https://csea.phhp.ufl.edu/media.html.

[174] M. Bradley and P. Lang. "The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual". In: *Technical Report B-3* (2007). URL: https://csea.phhp.ufl.edu/media.html.

[175] Beatrice Rammstedt and Oliver P. John. "Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German". In: *Journal of Research in Personality* 41.1 (Feb. 2007), pp. 203–212. DOI: 10.1016/j.jrp.2006.02.001. URL: https://doi.org/10.1016%2Fj.jrp.2006.02.001.

[176] David Watson, Lee Anna Clark, and Auke Tellegen. "Development and validation of brief measures of positive and negative affect: The PANAS scales." In: *Journal of Personality and Social Psychology* 54.6 (1988), pp. 1063–1070. DOI: 10.1037/0022-3514.54.6.1063. URL: https://doi.org/10.1037%2F0022-3514.54.6.1063.

[177] Alexandros Mileounis, Raymond H. Cuijpers, and Emilia I. Barakova. "Creating Robots with Personality: The Effect of Personality on Social Intelligence". In: *Artificial Computation in Biology and Medicine*. Springer International Publishing, 2015, pp. 119–132. DOI: 10.1007/978-3-319-18914-7_13. URL: https://doi.org/10.1007%2F978-3-319-18914-7_13.

[178] Edward W. P. Schafer. "Neural Adaptability: A Biological Determinant of Behavioral Intelligence". In: *International Journal of Neuroscience* 17.3 (Jan. 1982), pp. 183–191. DOI: 10.3109/00207458208985922. URL: https://doi.org/10.3109%2F00207458208985922.

[179] Francesco Corea. "AI Knowledge Map: How to Classify AI Technologies". In: *Studies in Big Data*. Springer International Publishing, Nov. 2018, pp. 25–29. DOI: 10.1007/978-3-030-04468-8_4. URL: https://doi.org/10.1007%2F978-3-030-04468-8_4.

[180] Samu Kumpulainen and Vagan Terziyan. "Artificial General Intelligence vs. Industry 4.0: Do They Need Each Other?" In: *Procedia Computer Science* 200 (2022), pp. 140–150. DOI: 10.1016/j.procs.2022.01.213. URL: https://doi.org/10.1016%2Fj.procs.2022.01.213.

[181] Jingtao Fan et al. "From Brain Science to Artificial Intelligence". In: *Engineering* 6.3 (2020), pp. 248–252. ISSN: 2095-8099. DOI: https://doi.org/10.1016/j.eng.2019.11.012. URL: https://www.sciencedirect.com/science/article/pii/S2095809920300035.

[182] Xiao Huang, Wei Wu, and Hong Qiao. "Connecting Model-Based and Model-Free Control With Emotion Modulation in Learning Systems". In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 51.8 (Aug. 2021), pp. 4624–4638. DOI: 10.1109/tsmc.2019.2933152. URL: https://doi.org/10.1109%2Ftsmc.2019.2933152.

[183] Haixu Yu and Pei Yang. "An Emotion-Based Approach to Reinforcement Learning Reward Design". In: *2019 IEEE 16th International Conference on Networking, Sensing and Control (ICNSC)*. IEEE, May 2019. DOI: 10.1109/icnsc.2019.8743211. URL: https://doi.org/10.1109%2Ficnsc.2019.8743211.

[184] Dung Nguyen et al. "Theory of mind with guilt aversion facilitates cooperative reinforcement learning". In: *Asian Conference on Machine Learning*. PMLR. 2020, pp. 33–48.

[185] Pedro Sequeira, Francisco S. Melo, and Ana Paiva. "Emergence of emotional appraisal signals in reinforcement learning agents". In: *Autonomous Agents and Multi-Agent Systems* 29.4 (Apr. 2014), pp. 537–568. DOI: 10.1007/s10458-014-9262-4. URL: https://doi.org/10.1007%2Fs10458-014-9262-4.

[186] Thomas M. Moerland, Joost Broekens, and Catholijn M. Jonker. "Emotion in reinforcement learning agents and robots: a survey". In: *Machine Learning* 107.2 (Aug. 2017), pp. 443–480. DOI: 10.1007/s10994-017-5666-0. URL: https://doi.org/10.1007%2Fs10994-017-5666-0.

[187] M. Kirtay et al. "Emergent emotion as a regulatory mechanism for a cognitive task implemented on the iCub robot". In: *Continual Unsupervised Sensorimotor Learning Workshop* (2018). URL: https://conferences.au.dk/fileadmin/conferences/2018/ICDL-EpiRob/Kirtay2018Emergent.pdf.

[188] Murat Kirtay et al. "Emotion as an emergent phenomenon of the neurocomputational energy regulation mechanism of a cognitive agent in a decision-making task". In: *Adaptive Behavior* 29.1 (Oct. 2019), pp. 55–71. DOI: 10.1177/1059712319880649.

[189] Xiao Huang, Wei Wu, and Hong Qiao. "Computational Modeling of Emotion-Motivated Decisions for Continuous Control of Mobile Robots". In: *IEEE Transactions on Cognitive and Developmental Systems* 13.1 (Mar. 2021), pp. 31–44. DOI: 10.1109/tcds.2019.2963545. URL: https://doi.org/10.1109%2Ftcds.2019.2963545.

[190] Cheng-Xiang Lu et al. "Using Emotions as Intrinsic Motivation to Accelerate Classic Reinforcement Learning". In: *2016 International Conference on Information System and Artificial Intelligence (ISAI)*. IEEE, June 2016. DOI: 10.1109/isai.2016.0077. URL: https://doi.org/10.1109%2Fisai.2016.0077.

[191] Joost Broekens and Laduona Dai. "A TDRL Model for the Emotion of Regret". In: *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, Sept. 2019. DOI: 10.1109/acii.2019.8925441. URL: https://doi.org/10.1109%2Facii.2019.8925441.

[192] Xiao Huang et al. "Brain-Inspired Motion Learning in Recurrent Neural Network With Emotion Modulation". In: *IEEE Transactions on Cognitive and Developmental Systems* 10.4 (Dec. 2018), pp. 1153–1164. DOI: 10.1109/tcds.2018.2843563. URL: https://doi.org/10.1109%2Ftcds.2018.2843563.

[193] Nikhil Churamani et al. "Affect-Driven Learning of Robot Behaviour for Collaborative Human-Robot Interactions". In: *Frontiers in Robotics and AI* 9 (Feb. 2022). DOI: 10.3389/frobt.2022.717193. URL: https://doi.org/10.3389%2Ffrobt.2022.717193.

[194] Guangliang Li et al. "Facial feedback for reinforcement learning: a case study and offline analysis using the TAMER framework". In: *Autonomous Agents and Multi-Agent Systems* 34.1 (Feb. 2020). DOI: 10.1007/s10458-020-09447-w. URL: https://doi.org/10.1007%2Fs10458-020-09447-w.

[195] Can Wang et al. "Unintentional Islanding Transition Control Strategy for Three-/Single-Phase Multimicrogrids Based on Artificial Emotional Reinforcement Learning". In: *IEEE Systems Journal* 15.4 (Dec. 2021), pp. 5464–5475. DOI: 10.1109/jsyst.2021.3074296. URL: https://doi.org/10.1109%2Fjsyst.2021.3074296.

[196] Pablo Barros et al. “Moody Learners - Explaining Competitive Behaviour of Reinforcement Learning Agents”. In: *2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)*. IEEE, Oct. 2020. DOI: 10.1109/icdl-epirob48136.2020.9278125. URL: https://doi.org/10.1109%2Ficdl-epirob48136.2020.9278125.

[197] Chie Hieida, Takato Horii, and Takayuki Nagai. “Emotion Differentiation based on Decision-Making in Emotion Model”. In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Aug. 2018. DOI: 10.1109/roman.2018.8525579. URL: https://doi.org/10.1109%2Froman.2018.8525579.

[198] Yizeng Han et al. “Dynamic Neural Networks: A Survey”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2021), pp. 1–1. DOI: 10.1109/tpami.2021.3117837. URL: https://doi.org/10.1109%2Ftpami.2021.3117837.

[199] Gao Huang et al. “Multi-Scale Dense Networks for Resource Efficient Image Classification”. In: *International Conference on Learning Representations*. 2018. URL: https://openreview.net/forum?id=Hk2aImxAb.

[200] Edward W Staley, Corban G Rivera, and Neil Joshi. *Triangular Dropout: Variable Network Width without Retraining*. 2022. URL: https://openreview.net/forum?id=B7abCaIiN_v.

[201] Amirkeivan Mohtashami, Martin Jaggi, and Sebastian Stich. “Masked Training of Neural Networks with Partial Gradients”. In: *Proceedings of The 25th International Conference on Artificial Intelligence and Statistics*. Vol. 151. Proceedings of Machine Learning Research. PMLR, Mar. 2022, pp. 5876–5890. URL: https://proceedings.mlr.press/v151/mohtashami22a.html.

[202] Yi-Lin Sung, Varun Nair, and Colin A Raffel. “Training Neural Networks with Fixed Sparse Masks”. In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato et al. Vol. 34. Curran Associates, Inc., 2021, pp. 24193–24205. URL: https://proceedings.neurips.cc/paper/2021/file/cb2653f548f8709598e8b5156738cc51-Paper.pdf.

[203] Ruoyu Li et al. “Adaptive Graph Convolutional Neural Networks”. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 32.1 (Apr. 2018). DOI: 10.1609/aaai.v32i1.11691. URL: https://doi.org/10.1609%2Faaai.v32i1.11691.

[204] Thomas Miconi et al. “Backpropamine: training self-modifying neural networks with differentiable neuromodulated plasticity”. In: *International Conference on Learning Representations*. 2019. URL: https://openreview.net/forum?id=r1lrAiA5Ym.

[205] Thomas Miconi, Kenneth Stanley, and Jeff Clune. “Differentiable plasticity: training plastic neural networks with backpropagation”. In: *Proceedings of the 35th International Conference on Machine Learning*. Vol. 80. Proceedings of Machine Learning Research. PMLR, July 2018, pp. 3559–3568. URL: https://proceedings.mlr.press/v80/miconi18a.html.

[206] Bernd Illing et al. "Local plasticity rules can learn deep representations using self-supervised contrastive predictions". In: *Advances in Neural Information Processing Systems*. Ed. by M. Ranzato et al. Vol. 34. Curran Associates, Inc., 2021, pp. 30365–30379. URL: https://proceedings.neurips.cc/paper/2021/file/feade1d2047977cd0cefdafc40175a99-Paper.pdf.

[207] Tong Yu and Hong Zhu. "Hyper-Parameter Optimization: A Review of Algorithms and Applications". In: 2003.05689 (2020).

[208] Takuya Isomura and Karl Friston. "Reverse-Engineering Neural Networks to Characterize Their Cost Functions". In: *Neural Computation* 32.11 (Nov. 2020), pp. 2085–2121. ISSN: 0899-7667. DOI: 10.1162/neco_a_01315. eprint: https://direct.mit.edu/neco/article-pdf/32/11/2085/1865423/neco_a_01315.pdf. URL: https://doi.org/10.1162/neco_a_01315.

[209] Karl Friston. "The free-energy principle: a unified brain theory?" In: *Nature Reviews Neuroscience* 11.2 (Jan. 2010), pp. 127–138. DOI: 10.1038/nrn2787. URL: https://doi.org/10.1038%2Fnrn2787.

[210] Boris Ginsburg, Igor Gitman, and Yang You. *Large Batch Training of Convolutional Networks with Layer-wise Adaptive Rate Scaling*. 2018. URL: https://openreview.net/forum?id=rJ4uaX2aW.

[211] Jesús Garciá Fernández, Enrique Hortal, and Siamak Mehrkanoon. "Towards biologically plausible learning in neural networks". In: *2021 IEEE Symposium Series on Computational Intelligence (SSCI)*. 2021, pp. 01–08. DOI: 10.1109/SSCI50451.2021.9659539.

[212] Victor Geadah et al. "Advantages of biologically-inspired adaptive neural activation in RNNs during learning". In: 2006.12253 (2020).

[213] Giorgia Dellaferrera et al. "Introducing principles of synaptic integration in the optimization of deep neural networks". In: *Nature Communications* 13.1 (Apr. 2022). DOI: 10.1038/s41467-022-29491-2. URL: https://doi.org/10.1038%2Fs41467-022-29491-2.

[214] Petar Radanliev and David De Roure. "Review of the state of the art in autonomous artificial intelligence". In: *AI Ethics* (June 2022). DOI: 10.1007/s43681-022-00176-2. URL: https://doi.org/10.1007%2Fs43681-022-00176-2.

[215] Bing Liu et al. "Self-Initiated Open World Learning for Autonomous AI Agents". In: 2110.11385 (2021).

[216] Tom Macpherson et al. "Natural and Artificial Intelligence: A brief introduction to the interplay between AI and neuroscience research". In: *Neural Networks* 144 (Dec. 2021), pp. 603–613. DOI: 10.1016/j.neunet.2021.09.018. URL: https://doi.org/10.1016%2Fj.neunet.2021.09.018.

[217] Yucan Chen et al. "How far is brain-inspired artificial intelligence away from brain?" In: *Front. Neurosci.* 16 (Dec. 2022). DOI: 10.3389/fnins.2022.1096737. URL: https://doi.org/10.3389%2Ffnins.2022.1096737.

[218] Jie Mei, Rouzbeh Meshkinnejad, and Yalda Mohsenzadeh. "Effects of neuromodulation inspired mechanisms on the performance of deep neural networks in a spatial learning task". In: *iScience* 26.2 (Feb. 2023), p. 106026. DOI: 10.1016/j.isci.2023.106026. URL: https://doi.org/10.1016%2Fj.isci.2023.106026.

[219] Kai Olav Ellefsen, Jean-Baptiste Mouret, and Jeff Clune. "Neural Modularity Helps Organisms Evolve to Learn New Skills without Forgetting Old Skills". In: *PLoS Computational Biology* 11.4 (Apr. 2015). Ed. by Josh C. Bongard, e1004128. DOI: 10.1371/journal.pcbi.1004128. URL: https://doi.org/10.1371%2Fjournal.pcbi.1004128.

[220] Arjun Magotra and Juntae Kim. "Neuromodulated Dopamine Plastic Networks for Heterogeneous Transfer Learning with Hebbian Principle". In: *Symmetry* 13.8 (July 2021), p. 1344. DOI: 10.3390/sym13081344. URL: https://doi.org/10.3390%2Fsym13081344.

[221] Nicolas Vecoven et al. "Introducing neuromodulation in deep neural networks to learn adaptive behaviours". In: *PLoS ONE* 15.1 (Jan. 2020). Ed. by William W Lytton, e0227922. DOI: 10.1371/journal.pone.0227922. URL: https://doi.org/10.1371%2Fjournal.pone.0227922.

[222] François Chollet et al. *Keras*. https://keras.io. 2015.

[223] Alex Krizhevsky, Vinod Nair, and Geoffrey Hinton. "CIFAR-10 (Canadian Institute for Advanced Research)". In: (2009). URL: http://www.cs.toronto.edu/~kriz/cifar.html.

[224] Susanne Nikolaus et al. "Investigating the Dopaminergic Synapse In Vivo. I. Molecular Imaging Studies in Humans". In: *Reviews in the Neurosciences* 18.6 (Jan. 2007). DOI: 10.1515/revneuro.2007.18.6.439. URL: https://doi.org/10.1515%2Frevneuro.2007.18.6.439.

[225] Denis Sheynikhovich, Satoru Otani, and Angelo Arleo. "Dopaminergic Control of Long-Term Depression/Long-Term Potentiation Threshold in Prefrontal Cortex". In: *J. Neurosci.* 33.34 (Aug. 2013), pp. 13914–13926. DOI: 10.1523/jneurosci.0466-13.2013. URL: https://doi.org/10.1523%2Fjneurosci.0466-13.2013.

[226] Ben-Hur Souto Neves et al. "On the role of the dopaminergic system in the memory deficits induced by maternal deprivation". In: *Neurobiology of Learning and Memory* 173 (Sept. 2020), p. 107272. DOI: 10.1016/j.nlm.2020.107272. URL: https://doi.org/10.1016%2Fj.nlm.2020.107272.

[227] H. Morikawa and C.A. Paladini. "Dynamic regulation of midbrain dopamine neuron activity: intrinsic, synaptic, and plasticity mechanisms". In: *Neuroscience* 198 (Dec. 2011), pp. 95–111. DOI: 10.1016/j.neuroscience.2011.08.023. URL: https://doi.org/10.1016%2Fj.neuroscience.2011.08.023.

[228] Gustavo Assunção, Miguel Castelo-Branco, and Paulo Menezes. "Self-mediated exploration in artificial intelligence inspired by cognitive psychology". In: *arXiv preprint arXiv:2302.06615* (2023).

[229] Jum C. Nunnally and L. Charles Lemond. "Exploratory Behavior and Human Development". In: *Advances in Child Development and Behavior Volume 8*. Elsevier, 1974, pp. 59–109. DOI: 10.1016/s0065-2407(08)60493-0. URL: https://doi.org/10.1016/s0065-2407(08)60493-0.

[230] Paula Kaanders et al. "Humans actively sample evidence to support prior beliefs". In: *eLife* 11 (Apr. 2022). DOI: 10.7554/elife.71768. URL: https://doi.org/10.7554%2Felife.71768.

[231] Lars Kunze et al. "Artificial Intelligence for Long-Term Robot Autonomy: A Survey". In: *IEEE Robotics and Automation Letters* 3.4 (2018), pp. 4023–4030. DOI: 10.1109/LRA.2018.2860628.

[232] Vieri Giuliano Santucci et al. "Editorial: Intrinsically Motivated Open-Ended Learning in Autonomous Robots". In: *Front. Neurorobot.* 13 (Jan. 2020). DOI: 10.3389/fnbot.2019.00115. URL: https://doi.org/10.3389%2Ffnbot.2019.00115.

[233] Delong Zhu et al. "Deep Reinforcement Learning Supervised Autonomous Exploration in Office Environments". In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. 2018, pp. 7548–7555. DOI: 10.1109/ICRA.2018.8463213.

[234] Pietro Mazzaglia et al. "Curiosity-Driven Exploration via Latent Bayesian Surprise". In: 2104.07495 (2022).

[235] Guido Schillaci et al. "Intrinsic motivation and episodic memories for robot exploration of high-dimensional sensory spaces". In: *Adaptive Behavior* 29.6 (June 2020), pp. 549–566. DOI: 10.1177/1059712320922916. URL: https://doi.org/10.1177%2F1059712320922916.

[236] Sébastien Forestier et al. "Intrinsically Motivated Goal Exploration Processes with Automatic Curriculum Learning". In: *Journal of Machine Learning Research* 23.152 (2022), pp. 1–41. URL: http://jmlr.org/papers/v23/21-0808.html.

[237] Michael R. W. Dawson. "Book Review - Computational neuroscience and cognitive modelling: A student's introduction to methods and procedures, By Britt Anderson, Thousand Oaks, CA: Sage Publications, 2014". In: *British Journal of Psychology* 105.3 (July 2014), pp. 436–438. DOI: 10.1111/bjop.12077. URL: https://doi.org/10.1111%2Fbjop.12077.

[238] M. Dawson et al. "Simple Artificial Neural Networks That Match Probability and Exploit and Explore When Confronting a Multiarmed Bandit". In: *IEEE Transactions on Neural Networks* 20.8 (Aug. 2009), pp. 1368–1371. DOI: 10.1109/tnn.2009.2025588. URL: https://doi.org/10.1109%2Ftnn.2009.2025588.

[239] Rebecca Alexander et al. "The neuroscience of positive emotions and affect: Implications for cultivating happiness and wellbeing". In: *Neuroscience & Biobehavioral Reviews* 121 (Feb. 2021), pp. 220–249. DOI: 10.1016/j.neubiorev.2020.12.002. URL: https://doi.org/10.1016%2Fj.neubiorev.2020.12.002.

[240] Mariann Oemisch et al. "Feature-specific prediction errors and surprise across macaque fronto-striatal circuits". In: *Nat Commun* 10.1 (Jan. 2019). DOI: 10.1038/s41467-018-08184-9. URL: https://doi.org/10.1038%2Fs41467-018-08184-9.

[241] David S. Stolz et al. "Internal control beliefs shape positive affect and associated neural dynamics during outcome valuation". In: *Nat Commun* 11.1 (Mar. 2020). DOI: 10.1038/s41467-020-14800-4. URL: https://doi.org/10.1038%2Fs41467-020-14800-4.

[242] Judith S. A. Asem, Felipe L. Schiffino, and Peter C. Holland. "Dorsolateral striatum is critical for the expression of surprise-induced enhancements in cue associability". In: *Eur J Neurosci* 42.5 (July 2015), pp. 2203–2213. DOI: 10.1111/ejn.13001. URL: https://doi.org/10.1111%2Fejn.13001.

[243] V. Srinivasa Chakravarthy and Pragathi Priyadharsini Balasubramani. "The Basal Ganglia System as an Engine for Exploration". In: *Computational Neuroscience Models of the Basal Ganglia*. Springer Singapore, 2018, pp. 59–96. DOI: 10.1007/978-981-10-8494-2_5. URL: https://doi.org/10.1007%2F978-981-10-8494-2_5.

[244] Tom Gilbertson and Douglas Steele. "Tonic dopamine, uncertainty and basal ganglia action selection". In: *Neuroscience* 466 (July 2021), pp. 109–124. DOI: 10.1016/j.neuroscience.2021.05.010. URL: https://doi.org/10.1016%2Fj.neuroscience.2021.05.010.

[245] V. Srinivasa Chakravarthy and Ahmed A. Moustafa. *Computational Neuroscience Models of the Basal Ganglia*. Springer Singapore, 2018. DOI: 10.1007/978-981-10-8494-2. URL: https://doi.org/10.1007%2F978-981-10-8494-2.

[246] Jaeeon Lee and Bernardo L. Sabatini. "Striatal indirect pathway mediates exploration via collicular competition". In: *Nature* 599.7886 (Nov. 2021), pp. 645–649. DOI: 10.1038/s41586-021-04055-4. URL: https://doi.org/10.1038%2Fs41586-021-04055-4.

[247] Sonja Utz and Nicole L. Muscanell. "Your Co-author Received 150 Citations: Pride, but Not Envy, Mediates the Effect of System-Generated Achievement Messages on Motivation". In: *Frontiers in Psychology* 9 (May 2018). DOI: 10.3389/fpsyg.2018.00628. URL: https://doi.org/10.3389%2Ffpsyg.2018.00628.

[248] Guido H. E. Gendolla. "Surprise in the Context of Achievement: The Role of Outcome Valence and Importance". In: *Motivation and Emotion* 21.2 (1997), pp. 165–193. DOI: 10.1023/a:1024486617134. URL: https://doi.org/10.1023%2Fa%3A1024486617134.

[249] Margaret Marshall and Jonathon Brown. "Emotional reactions to achievement outcomes: Is it really best to expect the worst?" In: *Cognition & Emotion* 20.1 (Jan. 2006), pp. 43–63. DOI: 10.1080/02699930500215116. URL: https://doi.org/10.1080%2F02699930500215116.

[250] Reinhard Pekrun. "Achievement emotions: A control-value theory perspective". In: *Emotions in Late Modernity*. Ed. by Roger Patulny et al. Routledge, Jan. 2019. DOI: 10.4324/9781351133319. URL: https://doi.org/10.4324%2F9781351133319.

[251] Li Deng. "The mnist database of handwritten digit images for machine learning research". In: *IEEE Signal Processing Magazine* 29.6 (2012), pp. 141–142.

[252] Mattis Hartwig and Achim Peters. "Cooperation and Social Rules Emerging From the Principle of Surprise Minimization". In: *Front. Psychol.* 11 (Jan. 2021). DOI: 10.3389/fpsyg.2020.606174. URL: https://doi.org/10.3389%2Ffpsyg.2020.606174.

[253] Jerrold H. Zar. *Spearman Rank Correlation: Overview*. Sept. 2014. DOI: 10.1002/9781118445112.stat05964. URL: https://doi.org/10.1002%2F9781118445112.stat05964.

[254] Xiao Yan Bi et al. "The influence of pride emotion on executive function: Evidence from ERP". In: *Brain and Behavior* 12.8 (July 2022). DOI: 10.1002/brb3.2678. URL: https://doi.org/10.1002%2Fbrb3.2678.

[255] Nathaniel J. Blanco et al. "Exploratory decision-making as a function of lifelong experience, not cognitive decline." In: *Journal of Experimental Psychology: General* 145.3 (Mar. 2016), pp. 284–297. DOI: 10.1037/xge0000133. URL: https://doi.org/10.1037%2Fxge0000133.

[256] Agnieszka Wykowska, Thierry Chaminade, and Gordon Cheng. "Embodied artificial agents for understanding human social cognition". In: *Phil. Trans. R. Soc. B* 371.1693 (May 2016), p. 20150375. DOI: 10.1098/rstb.2015.0375. URL: https://doi.org/10.1098%2Frstb.2015.0375.

[257] J. E. (Hans). Korteling et al. "Human versus Artificial Intelligence". In: *Frontiers in Artificial Intelligence* 4 (Mar. 2021). DOI: 10.3389/frai.2021.622364. URL: https://doi.org/10.3389%2Ffrai.2021.622364.

[258] Gustavo Assunção, Miguel Castelo-Branco, and Paulo Menezes. "Leveraging emotion-mediated exploration to adapt agent behavior". In: *2023 6th Experiment International Conference (exp.at'23)*. 2023.

[259] Gustavo Assunção et al. "Adapting Behavior and Persistence via Reinforcement and Self-Emotion Mediated Exploration in a Social Robot". In: *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, Aug. 2023. DOI: 10.1109/ro-man57019.2023.10309410. URL: https://doi.org/10.1109/ro-man57019.2023.10309410.

[260] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. "OpenFace: A general-purpose face recognition library with mobile applications". In: CMU School of Computer Science, 2016. DOI: CMU-CS-16-118.

[261] Rodolphe Gelin. "NAO". In: *Humanoid Robotics: A Reference*. Springer Netherlands, 2017, pp. 1–22. DOI: 10.1007/978-94-007-7194-9_14-1. URL: https://doi.org/10.1007%2F978-94-007-7194-9_14-1.

[262] Alexander Mordvintsev, Christopher Olah, and Mike Tyka. *DeepDream - a code example for visualizing Neural Networks.* Google Research, July 2015. URL: https://ai.googleblog.com/2015/07/deepdream-code-example-for-visualizing.html.

[263] Ilia Sucholutsky and Matthias Schonlau. "Soft-Label Dataset Distillation and Text Dataset Distillation". In: *2021 International Joint Conference on Neural Networks (IJCNN).* 2021, pp. 1–8. DOI: 10.1109/IJCNN52387.2021.9533769.

[264] Gustavo Assunção, Miguel Castelo-Branco, and Paulo Menezes. "ANNs Dream of Augmented Sheep: An Artificial Dreaming Algorithm". In: *Proceedings of the 2nd International Conference on Image Processing and Vision Engineering.* SCITEPRESS - Science and Technology Publications, 2022. DOI: 10.5220/0011055700003209.

[265] Xue Ying. "An Overview of Overfitting and its Solutions". In: *Journal of Physics: Conference Series* 1168 (Feb. 2019), p. 022022. DOI: 10.1088/1742-6596/1168/2/022022. URL: https://doi.org/10.1088/1742-6596/1168/2/022022.

[266] Susan Blackmore. *Consciousness: an introduction.* New York: Oxford University Press, 2012. ISBN: 0199739099.

[267] Christoph Adami. "What Do Robots Dream Of?" In: *Science* 314.5802 (Nov. 2006), pp. 1093–1094. DOI: 10.1126/science.1135929. URL: https://doi.org/10.1126%2Fscience.1135929.

[268] Deirdre Barrett. "The "committee of sleep": A study of dream incubation for problem solving." In: *Dreaming* 3.2 (June 1993), pp. 115–122. DOI: 10.1037/h0094375. URL: https://doi.org/10.1037%2Fh0094375.

[269] Sarah F. Schoch et al. "The effect of dream report collection and dream incorporation on memory consolidation during sleep". In: *J Sleep Res* 28.1 (Feb. 2019). DOI: 10.1111/jsr.12754. URL: https://doi.org/10.1111%2Fjsr.12754.

[270] Serena Scarpelli et al. "The Functional Role of Dreaming in Emotional Processes". In: *Front. Psychol.* 10 (Mar. 2019). DOI: 10.3389/fpsyg.2019.00459. URL: https://doi.org/10.3389%2Ffpsyg.2019.00459.

[271] Junyu Cao et al. "Unraveling why we sleep: Quantitative analysis reveals abrupt transition from neural reorganization to repair in early development". In: *Sci. Adv.* 6.38 (Sept. 2020). DOI: 10.1126/sciadv.aba0398. URL: https://doi.org/10.1126%2Fsciadv.aba0398.

[272] Erik Hoel. "The overfitted brain: Dreams evolved to assist generalization". In: *Patterns* 2.5 (May 2021), p. 100244. DOI: 10.1016/j.patter.2021.100244. URL: https://doi.org/10.1016%2Fj.patter.2021.100244.

[273] Rachit Dubey et al. “Investigating Human Priors for Playing Video Games”. In: *Proceedings of the 35th International Conference on Machine Learning.* Ed. by Jennifer Dy and Andreas Krause. Vol. 80. Proceedings of Machine Learning Research. Stockholmsmässan, Stockholm Sweden: PMLR, July 2018, pp. 1349–1357.

[274] Joshua M. Martin et al. “Structural differences between REM and non-REM dream reports assessed by graph analysis”. In: *PLoS ONE* 15.7 (July 2020). Ed. by Stavros I. Dimitriadis, e0228903. DOI: 10.1371/journal.pone.0228903. URL: https://doi.org/10.1371%2Fjournal.pone.0228903.

[275] Francesca Siclari et al. “Dreaming in NREM Sleep: A High-Density EEG Study of Slow Waves and Spindles”. In: *J. Neurosci.* 38.43 (Sept. 2018), pp. 9175–9185. DOI: 10.1523/jneurosci.0855-18.2018. URL: https://doi.org/10.1523%2Fjneurosci.0855-18.2018.

[276] Raffaele Manni. “Rapid eye movement sleep, non-rapid eye movement sleep, dreams, and hallucinations”. In: *Curr Psychiatry Rep* 7.3 (May 2005), pp. 196–200. DOI: 10.1007/s11920-005-0053-0. URL: https://doi.org/10.1007%2Fs11920-005-0053-0.

[277] Tomoyasu Horikawa and Yukiyasu Kamitani. “Hierarchical Neural Representation of Dreamed Objects Revealed by Brain Decoding with Deep Neural Network Features”. In: *Front. Comput. Neurosci.* 11 (Jan. 2017). DOI: 10.3389/fncom.2017.00004. URL: https://doi.org/10.3389%2Ffncom.2017.00004.

[278] A. Ranasinghe, R. Gayathri, and V. Priya. “Awareness of effects of sleep deprivation among college students”. In: *Drug Invention Today* (2018), pp. 1806–1809.

[279] David Foulkes. *Dreaming: A cognitive-psychological analysis.* Routledge, 2014.

[280] Alex Krizhevsky. “Learning Multiple Layers of Features from Tiny Images”. In: (2009), pp. 32–33. URL: https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.

[281] C. Daniel Batson. “These Things Called Empathy: Eight Related but Distinct Phenomena”. In: *The Social Neuroscience of Empathy.* The MIT Press, Mar. 2009, pp. 3–16. DOI: 10.7551/mitpress/9780262012973.003.0002. URL: https://doi.org/10.7551%2Fmitpress%2F9780262012973.003.0002.