Department of Computer Science
Series of Publications A
Report A-2022-3

Creative Systems, Agents and Societies: Theoretical Analysis Tools and Empirical Collaboration Studies

Simo Linkola

Doctoral thesis, to be presented for public examination with the permission of the Faculty of Science of the University of Helsinki, in Auditorium B123, Exactum Building, on the 19th of April, 2022 at 16 o'clock.

University of Helsinki
Finland

Supervisor
Hannu Toivonen, University of Helsinki, Finland

Pre-examiners
Simon Colton, Queen Mary University of London, UK
Oliver Bown, The University of New South Wales, Australia

Opponent
Dan Ventura, Brigham Young University, Utah, USA

Custos
Hannu Toivonen, University of Helsinki, Finland

Contact information

Department of Computer Science
P.O. Box 68 (Pietari Kalmin katu 5)
FI-00014 University of Helsinki
Finland

Email address: info@cs.helsinki.fi
URL: http://cs.helsinki.fi/
Telephone: +358 2941 911

ISSN 1238-8645 (print)
ISSN 2814-4031 (online)
ISBN 978-951-51-8025-4 (paperback)
ISBN 978-951-51-8026-1 (PDF)
Helsinki 2022
Unigrafia

Creative Systems, Agents and Societies: Theoretical Analysis Tools and Empirical Collaboration Studies

Simo Linkola

Department of Computer Science
P.O. Box 68, FI-00014 University of Helsinki, Finland
simo.linkola@helsinki.fi

PhD Thesis, Series of Publications A, Report A-2022-3
Helsinki, April 2022, 71+58 pages
ISSN 1238-8645 (print)
ISSN 2814-4031 (online)
ISBN 978-951-51-8025-4 (paperback)
ISBN 978-951-51-8026-1 (PDF)

Abstract

Creativity is a multi-faceted phenomenon that can be observed in diverse individuals and contexts, both natural and artificial. This thesis studies computational creativity, i.e. creativity in machines, which can be broadly categorised as a subfield of artificial intelligence. In particular, the thesis deals with three important perspectives on computational creativity: (1) identifying properties of creative individuals, (2) studying processes that lead to creative outcomes, and (3) observing and analysing social aspects of creativity, e.g. collaboration which may allow the individuals to create something together which they could not do alone.

One of the key interests in computational creativity is how computational entities may exhibit creativity in their own right, implying that the creative entities and their compositions, roles, processes and interactions are potentially different from those encountered in nature. This calls for theoretical analysis methods specifically tailored for artificial creative entities, and carefully controlled empirical experiments and simulations with them. We study both of these aspects. The analysis methods allow us to scrutinise exactly how creativity occurs in artificial entities by providing appropriate conceptual elements and vocabulary, while experiments enable us to test and confirm the effectiveness of different design decisions considering individual artificial creative entities and their interaction with each other.

We propose three novel, domain-general analysis tools for artificial creative entities, i.e. creative systems and creative agents, and collections of them, called creative societies. First, we distinguish several conceptual components relevant for metacreative systems, i.e. systems that can reflect and control their creative behaviour, and discuss how these components are interlinked and affect the system's creativity. Second, we merge elements from sequential decision making in intelligent agents, i.e. Markov Decision Processes, into formal creativity as search model called the Creative Systems Framework, providing a detailed account of various elements which compose the decision-making process of a creative agent. Third, we map elements from an eminent social creativity theory, the Systems View of Creativity, a.k.a. Domain-Individual-Field-Interaction model, into the elements of the Creative Systems Framework and show how creative societies may be analysed formally with it.

Each of the proposed analysis tools provides new ways to analyse creativity in artificial entities. The analysis of metacreative systems assumes an architectural point of view to creativity, which has not been previously addressed in detail. Deconstructing the decision-making process of a creative agent gives us additional means to discuss and understand why or how a creative agent selects certain actions. Lastly, the contributions to the creative societies are the first formal framework for their analysis.

We also investigate in two consecutive case studies collaborator selection in creative societies. In the first study, we focus on what kind of cues, e.g. selfish or altruistic, assist in choosing beneficial collaboration partners when all the agents can observe from their peers are the individually created end products. The second study allows the agents to adjust their aesthetic preferences during the simulations and inspects what emerges from society as a whole. We conclude that selfish cues seem to be more effective in choosing the collaboration partners in our settings and that the society exhibits distinct emergence depending on how much the agents are willing to change their aesthetic preferences.

Computing Reviews (2012) Categories and Subject Descriptors:

Computing methodologies → Artificial Intelligence → Philosophical/theoretical foundations of artificial intelligence
Computing methodologies → Modeling and Simulation → Model development and analysis

Computing methodologies → Modeling and Simulation →
Simulation types and techniques → Agent / discrete models
Software and its engineering → Software system structures

General Terms:
Theory, Experimentation, Creativity

Additional Key Words and Phrases:
computational creativity, creative system, creative agent, social creativity, self-awareness, metacreativity, creativity analysis, collaboration, role of self-awareness in metacreativity, formal analysis of creativity, collaboration simulations in abstract art, Creative Systems Framework

Acknowledgements

First and foremost, I would like to thank all the people that have made me the person I am today. Personal growth does not always happen through the most direct route, and I know I have made more than enough missteps in my life. For this thesis to be in your hands means that I have made it this far, and for that, I thank you all!

I sincerely thank Hannu Toivonen, Tomi Männistö, Anna Kantosalo, Niko Mäkitalo, Otto Hantula and Christian Guckelsberger for the collaboration on the papers included in the thesis. You made this thesis possible by your excellent cooperation skills and unyielding attitude towards research. I would also like to thank all the wonderful people I have had the pleasure to meet while working in the Department of Computer Science at the University of Helsinki. There are too many of you to be named here, but you know who you are. I express my gratitude to the Academy of Finland for supporting the research done in this thesis, and the Doctoral Programme of Computer Science for the guidance during my doctoral studies. I would like to show appreciation to my supervisor, professor Hannu Toivonen, for his exemplary instructions and leadership, it is easy to improve yourself when you can directly learn from competent professionals. Special thanks to the pre-examiners of this thesis, associate professor Oliver Bown and professor Simon Colton, for their invaluable comments on the thesis manuscript, and to my opponent, Dan Ventura, for agreeing to perform the most important role in the public defence of this thesis.

Personally, I would like to thank my mother and father, Pirjo and Antti, for all your support, and my brothers and sister, Kalle, Salla and Joonas, for being exactly who you are. I am also grateful for all my friends for providing me with emotional aid when I have needed it as well as being the mirror from which I can reflect on and learn about myself. Last, heartfelt thanks to Mikaela for your endless support during the thesis process.

Helsinki, March 2022
Simo Linkola

Contents

List of Publications

This thesis consists of an introductory part and the following original publications that are referred to as Papers I–V in the text. The publications are reprinted at the end of the thesis.

Metacreativity

Paper I: Simo Linkola, Anna Kantosalo, Tomi Männistö, and Hannu Toivonen. Aspects of Self-awareness: An Anatomy of Metacreative Systems. In *Proceedings of the Eight International Conference on Computational Creativity*, pp. 189–196, Atlanta, Georgia, USA, 2017.

> *Author's Contribution:* I actively participated in each stage and aspect of the research. I contributed to the the idea forming discussions regarding the paper and later took a major role in fleshing out the initial ideas to clear concepts. I also wrote a significant part of the paper.

Formal Frameworks for Creativity

Paper II: Simo Linkola, Christian Guckelsberger, and Anna Kantosalo. Action Selection in the Creative Systems Framework. In *Proceedings of the Eleventh International Conference on Computational Creativity*, pp. 303–310, Coimbra, Portugal, 2020.

> *Author's Contribution:* The paper started as a close collaboration with the co-authors. Towards the end of the process I took a leading role, sharpened some of the formulations and concepts, and wrote the majority of the paper.

Paper III: Simo Linkola and Anna Kantosalo. Extending the Creative Systems Framework for the Analysis of Creative Agent Societies. In *Proceedings of the Tenth International Conference on Computational Creativity*, pp. 204–211, Charlotte, NC, USA, 2019.

> *Author's Contribution:* I took a major role in all parts of the process and a leading role in writing the paper. That being said, the paper was a close collaboration with the co-author.

Empirical Collaboration Studies

Paper IV: Simo Linkola and Otto Hantula. On Collaborator Selection in Creative Agent Societies: An Evolutionary Art Case Study. In *Proceedings for the Seventh International Conference on Computational Intelligence in Music, Sound, Art and Design*, pp. 206–222, Parma, Italy, 2018.

> *Author's Contribution:* I had the leading role in guiding the research towards its target in discussions with Otto Hantula, where we fleshed out what we wanted to study and how the experiments would be arranged. Coding was done in close collaboration as was writing of the paper.

Paper V: Otto Hantula and Simo Linkola. Towards Goal-aware Collaboration in Artistic Agent Societies. In *Proceedings of the Ninth International Conference on Computational Creativity*, pp. 136–143, Salamanca, Spain, 2018.

> *Author's Contribution:* The research was a direct continuation from Paper IV. Otto Hantula wrote most of the code needed to run individual simulations, while I devised the experiment setups and implemented the majority of the analysis code. I also wrote parts of the paper.

Chapter 1

Introduction

Creativity is a ubiquitous and multi-faceted phenomenon. Glimpses of creativity, either as a single entity's behaviour or as a collective emergent behaviour of a community of entities, have been argued to be present not only in humans [92, 104] but also in other animals [54], in slime molds [1], and in artificial systems, e.g. computer software [8, 12, 25, 76].

1.1 The Topic of the Thesis

In this thesis, creativity is studied in the context of Computational Creativity (CC) [25]. We consider creativity as an abstract phenomenon, which may be observed both in artificial entities and in nature. *Our primary interest is in the creativity of an artificial, computational system or agent, and in the creativity in societies of such agents.* The thesis also considers *metacreativity*, the phenomenon where a system or an agent *reflects* and *controls* its own creative behaviour. Unsurprisingly, as the very idea of creativity is based on phenomena first observed in nature, many of the concepts considered in the thesis are inspired by – or at least the terms originate from – creativity found in living organisms.

In particular, the thesis deals with three distinct themes which all revolve around the design, analysis and empirical evaluation of computational systems, agents and societies exhibiting creativity in various ways.

The first theme [Paper I] identifies and isolates *aspects and components which are relevant for individual metacreative systems.* The work on this theme is theoretical in nature and draws from previous research in CC [e.g. 8, 12] and in software architectures [e.g. 55, 58, 61].

The second theme [Papers II and III] involves *formal, mathematical analysis frameworks which may be used to analyse both individual creative* *agents and creative agent societies.* Paper II considers individual creative agents and brings forth elements that can be used to analyse and compare action selection procedures in them. Paper III considers creative agent societies, i.e. sets of creative agents operating in the same environment and interacting with each other. It provides tools to analyse the society's dynamics, both how the agents differ from each other at particular times and how the society changes over time. Both of the papers extend a well-known formalism of creativity, the Creative Systems Framework (CSF) [115, 116, 118], which models creativity as search.

The third theme [Papers IV and V] covers *empirical research on efficient collaborator selection schemes between creative agents in situations where the agents do not share the same aesthetics* of what is considered desirable. Paper IV introduces the basic collaboration procedures for agents operating in the domain of abstract art and the first empirical simulations. Paper V adds complexity to the simulation setting by allowing the agents to adjust their aesthetics using a metacreative process during the simulation. In both of the papers, the technical implementations of individual agents are based on previous research on genetic programming [101] and generative art, while the interpretation of the results is done in the context of intentional collaboration and creative agent societies [96].

1.2 Creativity

The first studies of creativity in the field of psychology were done around the 1930s [92], and the current "standardised meaning" of the term creativity was first considered in the 1950s [92, 103]. The standard definition states that *"creativity requires both originality and effectiveness"* [92]. Often, this definition is personified, effectiveness is rephrased as usefulness, and originality is rephrased as novelty [8, 105]. This results in the following (standard) definition of creativity:

"Creativity is the ability to produce novel and useful artefacts"

where an artefact refers to the end product of the creative behaviour. In this thesis, words such as a *concept*, an *idea* and an *artefact* are all used to refer to the end product when not otherwise stated. Generally, the separation between non-materialised (concept, idea) and materialised (artefact) end products is not relevant for understanding the main points of the thesis. We will come back to this separation in Chapter 3 where we make a clear distinction between a creative agent and its operating environment.

In this thesis, we mainly conform to the above definition of creativity. However, as we will see later, this definition is still unable to capture all aspects of creative behaviour, both on an individual level and on a societal level (cf. Section 1.4). This has caused the dispute of whether creativity is still a contested concept to remain to this day [20].

Partly because of the ambiguity in the term's definition – which was facilitated by it originating from the separation of competing ideas such as freedom, imagination, genius, talent and individuality [2] – creativity was long denied the status as a field of psychology demanding its own attention [105]. The following are other notable reasons for this historical neglect. First, the origins of creativity research were in mysticism and spirituality deeming it as something untouchable or undefined [88, 105]. Second, some approaches to creativity argued that its study lacked the basis in psychological theory or verification through empirical research [105]. Third, other scholars viewed creativity as an extraordinary result of ordinary processes, raising the question of whether studying creativity as its own research field was necessary [105].

Despite past hindrances, the creativity research field has since then spawned multiple theories of (human) creativity. Kozbelt et al. [57] classify ten different types of creativity theories ranging from developmental approaches to problem solving and from economical theories to complex systems. Due to this collision of approaches and definitions, creativity is truly a multi-faceted phenomenon, where confluence theories are argued to be particularly apt approaches for its study [105]. Confluence theories state that multiple components must converge for creativity to occur, emphasising the need for multidisciplinary studies [105].

Today, creativity is noted as a worthy research subject and it is studied in fields such as history, psychology, social psychology, arts, fine arts, and artificial intelligence. Much of the current research interest stems from the fact that creativity is taken as one of the key factors which drive successful business and innovation [80, 105].[1] All in all, interaction, collaboration and cooperation are of special interest to creativity as *the creativity of an individual is influenced by other people, society and culture* [30, 80].

1.3 Computational Creativity (CC)

Computational Creativity (CC) studies creativity using machines. As a research field, it encompasses modelling, simulating and replicating (human) creativity in machines, and studies how machines may be creative in their own right, potentially exhibiting creative behaviours which differ from

[1] Innovation can be seen as a sister term for creativity which emphasises the organisational level [105].

those encountered in humans. Computational Creativity, in general, can be categorised as a subfield of Artificial Intelligence (AI) with a prominent multidisciplinary focus with connections to disciplines such as psychology, cognitive science, sociology, philosophy, literature, musicology and fine arts. However, CC has particular characteristics which distinguish it from artificial intelligence research as a whole. For example, CC often deals with open-ended problems where the best or optimal solution is either not defined or finding it does not render other solutions useless, the perception of creativity is highly subjective [8, 57, 103], and social aspects are pivotal for creativity as the end products of creative behaviour can only be understood fully in the context they were created [30, 35, 96]. Colton and Wiggins [25] characterise CC as

> *"the philosophy, science and engineering of computational systems which, by taking on particular responsibilities, exhibit behaviours that unbiased observers would deem to be creative".*

In this thesis, we adopt a slightly wider perspective to CC as the above characterisation obscures relevant aspects of creativity which take place when (1) a human and a computational system produce artefacts together or when (2) multiple computational systems work together. Namely, the point where a creativity support tool begins to *take responsibilities* – and thus co-creativity between a human and a computational system begins – is disputed and not well enough defined, and the inclusion of multiple independent creative systems operating in the same environment, e.g. creative societies, is not transparent in the characterisation. Our view is that research on computational creativity support tools and their development are included in the broad scope of CC, and we emphasise that it is necessary to also understand and study social aspects of creativity, on which we focus in Chapter 4.

Interest in machine creativity is at least as old as the Logic Theorist from the 1960s developed by Newell et al. [78]. However, the bulk of the CC field has developed during the last few decades, and the field is still evolving and growing today, partly due to the advent of robust generative AI methods. One of the pioneering studies in CC is Margaret Boden's [8] classification of three types of creative process: *combinational*, *exploratory* and *transformational creativity.* In combinational creativity, two previously known artefacts are combined into a new artefact. In exploratory creativity, a *search space* (which Boden dubs *conceptual space*) of possible solutions is searched for creative artefacts. Lastly, in transformational creativity, the search space is transformed in a way that allows reaching artefacts that could not be reached before. While the last two creativity types are more closely related to each other than either of them to the first one, Boden values transformational creativity the most as only it can produce something completely new. In this thesis, the main focus is on exploratory creativity, but transformational creativity is also touched upon. The thesis also considers a related term, *metacreativity*, the ability of a system to reflect and control its own creative process (see Section 2.5), which is seen as a major enabler of intentional transformational creativity.

Overall, Computational Creativity as a discipline encompasses all aspects of machine creativity, e.g. creativity support tools, such as digital audio workstations or image editing software, individual creative systems (operating in some specific domains), be it artificial or human, co-creativity between a human and a machine, and social aspects of creativity between multiple creative entities. In simplified terms, the goal of CC research may be either to implement (communities of) systems which may be seen as creative or support human creativity, to study human creativity using computational means, or to theorise and study how (communities of) computational systems may be creative in their own right. We refer to the first goal as applied CC and to the last two goals as theoretical CC, although we acknowledge that this categorisation does not do full justice to the eclectic spectrum of CC research.

In the applied study of CC, the goal is usually to design and implement a system that outputs end products (ideas or artefacts) in a specific domain where the end products are valuable (or useful) and at least in some aspects novel. That is, *the focus is on fulfilling the requirements for creativity.* There are two main approaches towards this goal: engineering and cognitive [82], however, they are not mutually exclusive. In the engineering approach, the focus is on the useful and efficient usage of processes and resources, such as machine learning, knowledge bases, or other expertise directly embedded in the system by its developer. In the cognitive approach, the main focus is on computational modelling of (some aspects of) cognitive functionality present in animals (typically humans) and investigating how this functionality allows or enhances creativity in a specific system.

Creative systems have been designed and implemented, e.g. in the following domains: mathematical theories [16, 18], dance choreographies [13, 29, 85], music [32, 33, 43, 47, 79], code generation [22, 27], image generation [21, 73, 101, 121], games [42, 62], figurative language [3, 120] and story generation [83, 84]. For a more comprehensive overview, we refer the reader to a quite recent survey on domains present in CC [70]. In this thesis, Paper IV and Paper V have elements of applied study in them.

In the theoretical CC, creativity is studied as it is or as it could be, potentially without a specific application domain or a class of systems in mind. Theoretical research varies from philosophical studies, e.g. drawing parallels between animal and machine creativity, to mathematical formalisms of certain aspects of creativity or phenomena related to it. That is, *the focus is on understanding the underlying mechanisms of creativity or developing new mechanisms.*

Goals and approaches of individual theoretical studies are diverse: e.g. to deepen our understanding of what creativity and computational creativity are [8, 12, 76, 96], to argue about a disputed aspect of (computational) creativity [71], to conceptualise and assess development of creative systems [23, 110, 111], or to provide tools for analysis, evaluation and comparison of creative systems [17, 50, 51, 89, 115]. In this thesis, Papers I–III are purely theoretical in nature, while Paper IV and Paper V study theory-oriented aspects of social creativity in an applied context.

1.4 Four Perspectives on Creativity

One of the categorisations by which (computational) creativity may be inspected, analysed and evaluated is the four perspectives on creativity, or simply the Four Ps [51, 74, 88]. The Four Ps addresses problems arising from the multi-faceted nature of creativity by separating its analysis into four different perspectives. Each of the perspectives may be adopted when elaborating and arguing about creativity as an abstract phenomenon or creativity of a particular entity (or community of entities). The perspectives are *Product*, *Producer*, *Process* and *Press*.

The perspectives have different requirements for, or interests in, an entity's creativity. Taking into consideration which perspectives are currently in focus assists in keeping creativity analysis relevant. In this thesis, we use the Four Ps as an overarching conceptualisation of the vast creative phenomenon, and, at times, refer to the perspectives to guide the reader's attention to relevant aspects of the subject at hand.

Next, we introduce each of the Four Ps and describe on a general level how they are present in the thesis.

1.4.1 Product

The Product perspective deals with the end product of a creative act or behaviour, i.e. artefact or idea. At a minimum, it attends to what kind of end products are seen as valuable or useful and what kind of novelty is observed in the product [8, 92]. That is, the standard definition of creativity directly considers the Product perspective. Naturally, both value and novelty can be strongly domain-dependent and, in some cases, non-experts may value different end products when compared to experts of the field [45, 59]. Notably, the end product of a creative behaviour may be a creative system itself (see Chapter 2 and transformational creativity).

The end products may also have other properties, which are not properly captured in the usefulness (or value) and novelty division. Such properties include aesthetics [6] and typicality [89], and also general requirements of or restrictions on the end products. For example, a piece of music may be seen as valuable, aesthetically pleasing and novel (i.e. not pastiche), but it may not fit into a certain genre of music. Thus, the product was a failure if the goal was to produce music in that particular genre. The Product perspective is present in Chapter 2 where we discuss creative systems and their analysis, and in Chapter 4 where we consider social aspects of creativity.

1.4.2 Producer

The Producer perspective covers the traits, characteristics and abilities of the entity, artificial system or natural, which exhibits creative behaviour. In humans, this means, e.g. personality, intellect, temperament, physique, attitudes and self-concept [88]. The Producer perspective is seen indirectly in the standard definition as it considers the production of creative end products as an ability of the producer, however, it does not state anything about the particular characteristics that are required or desired from the producer to be creative.

In CC, the cognitive approach typically conceptualises characteristics that are inspired by those encountered in humans or are oriented to understand creative entities from a top-down perspective. For example, Colton's [17] creative tripod claims that appreciation, skill and imagination are requirements for a creative entity (see Section 2.2.4). The engineering approach concentrates on characteristics on a technical or mathematical level with no aim to draw parallels between the characteristics and humans. Typical for the engineering approach is to focus on the (finer) mechanical or computational details of the system, such as the physical composition of a particular robot, software stack used to produce certain functionality or the architecture of the system. In this thesis, the Producer perspective is most apparent in Chapter 2, where we cover creative systems and their architectural concerns.

1.4.3 Process

The Process perspective inspects what the creative entity does when it exhibits creative behaviour. In humans, this would mean all the continuous psychological processes such as perception, thinking and learning [88], or more long-lasting processes such as the succession of stages in creative processes: preparation, incubation, insight and verification [74, 94]. In CC, processes may be any computational processes, which either aim to mimic processes in animals (cognitive approach) or are only found in machines (engineering approach). These processes may be statistical methods such as machine learning, data mining and generative models [107], and other computational processes, e.g. which constitute a reflection and control loop in a metacreative system.

The Process perspective also includes so-called micro-level processes. In CC this level includes computational procedures and routines which take care of smaller tasks that are utilised in distinct parts of the whole creative process. In general, these tasks can be anything, such as web scraping, regular expression matching and pathfinding, but the emphasis is placed on tasks typical – or even special – to creativity, e.g. computation of the aesthetic quality of an image.

The perspective is inspected more closely in Chapter 3, which focuses on the action selection process of creative agents and our contributions to it [Paper II]. The perspective is also present in Chapter 4 as collaboration and dynamics of an evolving agent society can be seen as (distributed) creative processes [Papers III–V].

There is a close relationship between the Process and the Producer perspectives. Especially in CC, if the system employs particular processes during its creative behaviour, they may also be seen to enable or facilitate certain abilities or characteristics of the system. For example, many different machine learning methods may give a system an ability to learn (Producer perspective), investigating how the learning exactly takes place is then analysis from the Process perspective.

1.4.4 Press

The Press perspective focuses on the context where the creative behaviour is performed and the relationships between the creative entity and other entities, e.g. other humans or artificial agents in society. In other words, it emphasises creativity as a situated activity that, to be understood fully, has to be observed and analysed in its context. It is focal to creativity as individuals are always situated in some society and the adopted culture influences their creativity [30], and interaction between individuals is present in many creative behaviours, e.g. collaboration [80]. For individual creative systems, the perspective is, perhaps somewhat counter-intuitively, topical as *creative autonomy will emerge only from interactions with others* [49].

The perspective covers both smaller-scale social aspects, such as creativity within a group or creativity between multiple entities, and larger-scale social structures, such as creativity within a specific culture. The perspective also aims to answer questions such as “Who evaluates creativity, and why?” or “Whose evaluation matters for an artefact to be accepted to a cultural domain?”, bringing into focus the relationships between the created end product, its creator and its assessors.

The Press perspective is implicitly present in the standard definition of creativity as the assessment of value and novelty are typically culturally influenced. On the other hand, the characterisation of CC, introduced in Section 1.3, touches the Press perspective by pointing out that the evaluation of creative behaviour should be done by an *unbiased observer* [25]. By coining the term *unbiased observer*, the characterisation aims to make it clear what is meant by CC. It does not claim that every assessor of every CC system should or needs to be unbiased. Specifically, the characterisation brings forth the problematic position of CC: human assessors may either underrate the observed creativity of an artefact if they know that it is made by a computer [76, 77] or, in other circumstances, they may assess it as more creative [28].

In this thesis, discussion of general social factors of creativity is based on the Systems View of Creativity [30], also known as the Domain-Individual-Field-Interaction model (DIFI model) [35], which is a theory of human creativity emphasising interaction and cultural aspects. The Press perspective, the DIFI model, and our contributions to social creativity [Papers III–V] are covered in Chapter 4.

1.5 Research Tasks

The thesis aims to provide outcomes to three *Research Tasks* (RTs), which are aligned with the three themes introduced at the beginning of this thesis. *The general context for all of the research tasks is creativity in artificial systems.* Even though some of the research task outcomes may be applicable also to human creativity, we do not consider those aspects in this thesis.

The individual research tasks are described below, accompanied by their specific research contexts and the perspectives on creativity that are emphasised in the research tasks.

Research Task I:

Design a tool to describe and characterise artificial metacreative systems. This research task is conducted in the context of software architectures and the aim is to identify elements that contribute to the metacreative capabilities of a creative system in qualitatively distinct ways. The research task is most closely related to the Producer perspective, as the focus is on higher-level conceptual capabilities and not on the implementation details of metacreativity.

Research Task II:

Design formal analysis tools for (a) the artefact creation process of individual creative agents and (b) processes in creative societies. The context for this research task is in the descriptive analysis of artificial creative systems and collectives of them. Analysing processes of individual creative agents (a) should take into account the specifics of the artefact creation process and be aligned with general decision-making processes in intelligent agents. The analysis of processes in creative societies (b) should, in turn, be able to describe meaningful situations which may occur in creative societies and be grounded on existing frameworks for social creativity. Both points, (a) and (b), naturally deal with the Process perspective, while the latter point (b) also considers the Press perspective.

Research Task III:

Design and implement a society of creative agents to study different collaborator selection procedures and observe phenomena emerging from those procedures. The context of this research task is social creativity and collaboration of independent creative agents. The implementation is done in the domain of abstract art, for which agents utilise genetic programming. The research task concentrates on outcomes dealing with the Process and the Press perspectives, while the implementation of the agent society needs to also consider the Product and the Producer perspectives, i.e. how the agents evaluate artefacts and what their capabilities are.

1.6 Contributions of the Thesis

The scientific contributions of the thesis are in the five original publications, Papers I–V, each contributing to the three Research Tasks of the thesis (see Section 1.5).

Paper I: Aspects of Self-Awareness: An Anatomy of Metacreative Systems

> In this paper, we propose concepts that are relevant for metacreativity and show how these concepts can provide meaningful characterisations of metacreative systems. Further, the paper provides a blueprint architecture for how these concepts typically fit together in a metacreative system. Although the base of this research draws from previous research in self-aware systems [61], this is the first paper that directly considers concepts of self-aware systems in the context of computational creativity and metacreative systems. (Research Task I)

Paper II: Action Selection in the Creative Systems Framework

> In this paper, we identify and isolate elements that play a role in the action selection process of a creative agent. The paper extends Wiggins' formal framework of Boden's exploratory (and transformational) creativity (also dubbed the Creative Systems Framework (CSF))[115, 116, 118]. The extension makes the action selection process explicit within the CSF and adds detail to the agent's goal in the CSF by separating a concept, an agent's inner representation of an idea, and an artefact, the concept's external materialised expression. (Research Task II)

Paper III: Extending the Creative Systems Framework for the Analysis of Creative Agent Societies

> In this paper, we propose the first mathematical framework directly developed for the analysis of social creativity. The paper embeds a well-known social creativity model, the Systems View of Creativity (DIFI model) [30, 35], into the Creative Systems Framework (CSF) [115, 116, 118]. Particularly, we consider multiple agents which exhibit exploratory creativity in the CSF. We show how the consequences of the agents' aggregated societal decision making may be seen as elements of the CSF on

another level of abstraction, and how the elements of the DIFI model can be seen on this abstraction level. (Research Task II)

Paper IV: On Collaborator Selection in Creative Agent Societies: An Evolutionary Art Case Study

In this paper, we study collaboration processes that are beneficial for heterogeneous agents with a goal to both create abstract art and collaborate with each other. The paper introduces the first collaboration model for individual agents using genetic programming [101], and proposes straightforward ways of what kind of cues from other agents' creations imply beneficial collaboration results. Using empirical simulations, the paper shows that some of the cues result in more fruitful collaborations. (Research Task III)

Paper V: Towards Goal-aware Collaboration in Artistic Agent Societies

In this paper, we extend the setting from Paper IV into a dynamic system where the agents can change their aesthetic preferences of the abstract art during the simulation. This results in a more complex experimental setting and we analyse the dynamics of the simulation as well as the collaboration results. By analysing the results, we argue that if the agents are inclined to change their aesthetic preferences small steps at a time, it results in a more stable society where the agents can use their accumulated expertise on their aesthetic settings better. (Research Task III)

1.7 The Structure of the Thesis

In this chapter, we introduced the topic of this thesis, gave an overview of creativity research and its origins, and introduced CC as a multidisciplinary field. We considered the four perspectives to creativity, elaborating how they are present in CC and in this thesis. Further, we presented the research tasks of the thesis and the contributions of the original articles included in the thesis. The rest of this thesis is organised as follows.

In Chapter 2 we introduce creative systems, i.e., systems that have a goal in producing valuable and novel artefacts. We consider both, analysis of such systems and how they are built focusing on general architectural components and patterns. The chapter provides background for the next chapters and covers Paper I and Research Task I.

In Chapter 3 we cover creative agents, a specific type of creative system which emphasises the separation between the system (agent) and its environment. The *actions* the agent takes in its environment are in focus when discussing creative agents, i.e. how the creative process of the agent unfolds as it makes observations of its environment and acts based on those observations. Paper II and the individual agent case of Research Task II are covered in this chapter.

In Chapter 4 we widen our perspective from single agents to multiple agents operating in the same environment, shifting the focus from all actions to *interactions* between agents. Our point of view to social creativity originates from studies of human creativity, but we concentrate on how it fits into the context of CC. Moreover, we cover collaboration between creative agents. Papers III–V, the society case of Research Task II, and Research Task III are covered in this chapter.

In Chapter 5 we discuss the outcomes of the research tasks defined in Section 1.5. The introductory part of the thesis ends in conclusions in Chapter 6, after which the original articles included in the thesis are presented.

Chapter 2

Creative Systems

This chapter covers the main characteristics of *creative systems*, their analysis and their general architectural concerns. By a *system*, we mean a set of mechanisms and functions which are tightly grouped to produce some desired behaviour, such as a piece of software. That is, a system is a single entity, but it contains multiple components which interact with each other. In particular, we do not consider in this chapter systems of systems, such as societies of creative entities.

The chapter presents our work on metacreative systems and their architectures [Paper I] and provides background for the following chapters. Principally, the chapter handles creative systems as abstract entities with no specific target system in mind but also gives a few examples in places where they are seen as helpful for grasping abstract concepts. The chapter focuses on the Producer perspective of creativity.

First, we formulate the basics of what we mean by a creative system. Then, we summarise some analysis tools previously proposed in Computational Creativity (CC). We cover one of the analysis tools in depth, the Creative Systems Framework (CSF) [115, 116], as we extend it directly in two different ways in Chapters 3 [Paper II] and 4 [Paper III]. Lastly, we consider general architectural aspects of (meta)creative systems, i.e. what kind of conceptual components or layouts are useful for creative systems, and our contributions to them [Paper I]. Our contributions can then, in turn, be used as analytical tools for metacreative systems.

2.1 Definition of a Creative System

There are multiple possible ways to define creative systems, each of which emphasises some of the four perspectives on creativity. We adopt the standard definition of creativity [8, 92] aligning it with the characterisation of CC [25] by defining creative systems in the following way.

Definition 1. *A creative system is capable of generating new and valuable ideas or artefacts.*

The above definition is common in CC, especially in applied research where the main focus is on the end products and fulfilling the requirements of creativity. Similar to the standard definition of creativity, Definition 1 does not consider *how* the generation is done nor *who* assesses (and by which standards) the novelty and value of the end products. As such, the definition omits the Process and the Press perspectives. However, this does not mean that the Process or the Press perspectives would not be interesting or worthwhile to analyse in creative systems, only that we focus on systems that satisfy Definition 1. The typical minimal assumption for the Process perspective is that the generation process is not merely random [see, e.g. 49, 110], e.g. picking a random painting from a huge database and showing it to an assessor which has not seen the painting before, as little creativity can be attributed to such a process. The assessor can be the system itself, another computational system or a human, e.g. the system's designer or an outsider. In the majority of the cases, the final assessment is done by humans. For any form of meaningful creativity, however, the system has to at least partially assess its own ideas or artefacts during or after the generation process [110].

The definition also specifies that the system has to have the overall *capability* for generating novel and valuable end products. It does not directly state if the system ever does so in practice, although a system that cannot generate a single novel and valuable end product is typically not considered creative. This remark is noteworthy as the ability of a creative system to generate apt end products may depend on many factors such as its initialisation, which may require some input parameters from the user or include some randomness, its current state, and its history. Thus, even a system that, when randomly initialised, generates a novel and valuable end product on every hundredth run is considered creative. Naturally, a system that does so on every run could be argued to reach its supposed creativity goal better and, in some aspects, be more creative.

2.2 Analysis of Creative Systems

Analysis of creative systems aims to provide tools that can be applied to either individual creative systems or their collectives to extract useful information regarding them. Analysis of creative systems is central for this thesis as Papers I–III all propose conceptualisations and formalisms which can be used in analysing creativity.

The analysis tools may, e.g., describe required properties for creative systems, characterise relevant phenomena which may occur within creative systems or provide guidelines for quantifying particular dimensions of creative systems. Often, the analysis results can be used to evaluate creative systems or even compare them, albeit arguments have been raised analogous to human creativity research of whether CC is domain general [71, 86] making the comparison between creative systems operating in different domains challenging. On the other hand, as creativity is a multi-faceted phenomenon, it is commonly not considered appropriate to try to fit creative systems into a one-dimensional line from less creative to more creative. Instead, we should ask *where* and *how* each system exhibits creative behaviour [8], possibly doing so separately for all the four Ps.

Next, we summarise some analysis tools, frameworks, properties and dimensions previously proposed in CC. The intention is not to provide an all-encompassing overview of all the proposed methods but rather a selection that is broad enough to understand the field in general. For a more comprehensive overview of analysis and evaluation methods in CC, we refer the reader to the recent survey by Lamb et al. [60].

2.2.1 Boden's Three Types of Creativity

Perhaps the best-known classification of creative systems is Boden's [8] three types of the creative process: *combinational*, *exploratory* and *transformational* creativity. Combinational creativity combines two (or more) concepts in a new way resulting in a novel and valuable end product. Exploratory creativity models creativity as search in some known *conceptual space*, which can be any structured style of thought. As the search unfolds, the system may come up with creative concepts in different areas of the conceptual space. Examples of conceptual spaces are buildings in a certain architectural style, mathematical proofs or jazz music. The conceptual spaces can also be hierarchically nested, e.g. the conceptual space of paintings may be divided into paintings in certain styles and the resulting subspaces may be further divided. Transformational creativity aims to transform one of the conceptual spaces, providing a way for the system to reach concepts that it could not produce before the transformation(s). Boden [7] argues that the two latter types of creativity are closer to each other than either of them are to the first one and that transformational creativity is the only way to come up with something completely original.

2.2.2 Subjectivity

One of the main categorisation dimensions of creativity is its *subjectivity* or magnitude [57]. This dimension includes terms P-creativity (where P stands for personal or psychological) and H-creativity (where H stands for historical) [8] and smaller c (often more subjective) and Larger C creativity (more objective) [103]. For the latter term pair, a typical dichotomy is given as little-c (everyday creativity) and Big-C (eminent creativity) [57].

Subjectivity is closely related to the Press perspective; it focuses on who assesses creativity and separates *creative experience* from the end product. As such, subjectivity should not be taken as a binary choice between P-creativity and H-creativity (or little-c and Big-C), but rather as a complex interplay between the context where creative behaviour occurs and which entities are evaluating it.

2.2.3 Ritchie's Empirical Criteria

Ritchie [89] provides a set of empirical criteria that can be used to analyse or evaluate a computational system's creativity from the Product perspective. Ritchie considers systems that generate objects in a class of artefacts. The class should not be defined by the workings of the system and the humans should be able to assess the 'quality' (value) of those artefacts. Moreover, the class inclusion may be 'fuzzy' or subjective. Ritchie also defines a set of basic items which are available for the system (i.e. not generated by the system but, e.g. given by the programmer or available through web search), which he calls an *inspiring set*. Ritchie, then, proceeds to provide suggestions of criteria for the artefacts generated by the system. The criteria are based on novelty, quality, typicality (how typical the artefact is as a member of its class) and how the artefacts relate to the inspiring set.

2.2.4 Colton's Creative Tripod

The creative tripod characterises that a computational system can be perceived to be creative only if it has three required behaviours or abilities: *skill*, *appreciation* and *imagination* [17]. Each of the tripod's legs is split into three further parts which may be perceived as contributing creatively: the consumer, the programmer and the system itself. The main argument is that if the system is perceived to be skilful, appreciative and imaginative, *regardless of the consumer's and programmer's behaviour*, the system can be said to be creative. As the creative tripod focuses on the system's abilities it provides analysis from the Producer perspective. Nonetheless, the creative tripod reminds us that also the programmer is an actor in computational creativity, which is missing from the four Ps.

2.2.5 FACE and IDEA Descriptive Models

The FACE descriptive model considers creative acts as tuples of generative acts [19, 81]. It distinguishes between Concepts, which are programs, Expressions of a concept, i.e. an (input, output) pair of a program, Aesthetics measures, i.e. functions which take an expression of a concept and output a real value, and Framings, i.e. pieces of natural language readable by humans which aim to explain (part of) the generation process. The model further separates these four object types from methods of generating them. For example, a system may have an aesthetic measure, but it may also have a way to generate new aesthetic measures by which the system may acquire transformational creativity (see Boden's [8] three types of creativity).

The IDEA descriptive model [19, 81] describes a creative system's *potential impact.* IDEA stands for Iterative Development-Execution-Appreciation cycle, where a system is developed further and its behaviour is exposed to an audience in iterative cycles. The IDEA model suggests the following stages of engineering for creative systems: the developmental stage, fine tuned stage, re-invention stage, discovery stage, disruption stage and disorientation stage, where the stages proceed in the order of how similar the creative acts taken by the system are to the known ones. Colton et al. note, that creative systems are unlikely to have much impact outside of discovery and disruption stages [19].

The FACE and IDEA models can both be seen to provide analysis capabilities for multiple perspectives. First, the overall capability for different creative acts in FACE and the system's current stage of engineering in IDEA can be seen to be abilities, i.e. related to the Producer perspective. Second, software's evolution during its development can be described both in FACE and IDEA models, placing them as analysis tools for the meta-level process of software development. Third, the creative acts in FACE can be used to classify single artefact creation processes (Process perspective).

2.2.6 Computational Creativity Continuum

The computational creativity continuum (CC-Continuum) [82] is an analysis dimension that is tailored to describe the approach the designers of a creative system are using. It can be understood as a meta-level analysis tool, which aids in focusing the analysis on the right elements or properties of the system. At the other end of the dimension is the *engineering-*

mathematical approach and in the other end the *cognitive-social* approach, but systems can also be developed with a mixed approach.

The engineering-mathematical approach is characterised by the usage of techniques that have been developed for other purposes than creative systems, such as optimisation, probabilistic machine learning and general problem-solving techniques. While developing a system using this approach, much of the resources are spent on figuring out how to make the end products more appealing.

The cognitive-social approach, on the other hand, uses human (or animal) behaviour as a basis to develop computational models for creative systems. The goal may be, e.g., to develop a model of how humans come up with new ideas, to study how the quality of a piece of music is assessed by humans and replicate it as a computational method, or to study how social contexts affect the creative processes of individual entities. Thus, the main tasks of this approach include how to represent cognitive, social or cultural behaviours in computers.

The continuum is a cross-cutting dimension that is not fully contained in any of the four Ps. Nonetheless, from the above characterisation of the dimensions, we can derive that the engineering-mathematical approach emphasises the Product perspective, while the cognitive-social approach emphasises the other three Ps. However, even research focusing on the cognitive-social approach typically intends to create apt artefacts.

2.2.7 SPECS

The Standardised Procedure for Evaluating Creative Systems (SPECS) proposed by Jordanous [50] is a meta-level analysis and evaluation tool for creative systems. It defines a procedure that offers an unambiguous analysis of a system's creativity. As it is a meta-level tool, it does not itself provide any criteria for creativity. SPECS consists of the following steps:

1. Identify a definition of creativity that your system should satisfy to be considered creative,

2. Using Step 1, clearly state what standards you use to evaluate the creativity of your system, and

3. Test your creative system against the standards stated in Step 2 and report the results,

where in Step 1, the goal is both to determine general criteria for creativity and to identify any specific criteria or particularly prominent aspects of creativity for the domain in which the system produces artefacts.

The main claim of SPECS is that after utilising and reporting the choices made in each of the three steps, the analysis and evaluation of creativity should be transparent for the reader.

2.3 The Creative Systems Framework (CSF)

The Creative Systems Framework (CSF) introduced by Wiggins [115, 116, 118] is one of the frameworks or tools by which creative systems may be analysed. The CSF models creativity as search and, as stated originally by Wiggins [115, 116], it is a formal account of Boden's descriptive hierarchy of creativity (see Section 2.2.1). However, it does not explicitly include combinational creativity, which is one of Boden's three creativity process types. Nonetheless, combinational creativity can be modelled also as search in space that has a specific kind of structure.

The CSF allows for formal, mathematical discussion of creative systems and the description of relevant phenomena within them, which particularly concerns the mechanisms of *exploratory* and *transformational creativity*. In the analysis implementations, its current focus has been on the Producer perspective, as its main applications characterise different abilities or capabilities a creative system may have. Covering both exploratory and transformational creativity, the CSF can also be used to formally categorise the overall Boden's process type of a creative system, however, its capabilities are limited in describing an ongoing creative process of a system. For example, it treats the function which selects where the system moves next as a black box, and it does not have any defined stopping criteria.

The CSF is used as the basis in several studies in CC [e.g. 4, 41, 52]. In this thesis, the exploratory part of the CSF is directly extended in two different ways in Paper II and Paper III, which we cover in Chapter 3 and Chapter 4, respectively.

Next, we first introduce the *object level* part of the CSF which deals with exploratory creativity, and then the *meta-level* part of the CSF, which adjusts the object level and allows transformational creativity. We focus on the object level and give only a brief explanation of the meta-level CSF, as the contributions of this thesis do not consider it in detail.

2.3.1 Exploratory Creativity

Exploratory creativity consists of discovering novel and valuable concepts within a known conceptual space [8]. It is similar to the classic AI search paradigm with the prominent distinction of explicitly defining the conceptual space of (currently) acceptable solutions. In a typical AI search solution, the system is engineered in a way that it cannot escape the preferred conceptual space. An explicit definition of the conceptual space allows a system to make informed decisions regarding it, possibly provoking transformational creativity, which we cover in the next section.

The CSF defines exploratory creativity as a septuple

$$\langle \mathcal{U}, \mathcal{L}, [\![.]\!], \langle\langle ., ., . \rangle\rangle, \mathcal{R}, \mathcal{T}, \mathcal{E} \rangle, \tag{2.1}$$

where the individual elements of the septuple are characterised in Table 2.1. Below, we explain each element and its role in the system's exploratory creativity, concentrating on the elements more relevant to this thesis.

The *universe* $\mathcal{U}$ is a multidimensional (possibly infinite-dimensional) space capable of representing *anything*. All possible distinct concepts $c \in \mathcal{U}$ are distinct points in $\mathcal{U}$. The empty concept $\top$ is also part of the universe, $\top \in \mathcal{U}$. The language $\mathcal{L}$ (the set of all words over some vocabulary $\mathcal{V}$, $\mathcal{L} = \mathcal{V}^*$) is needed for completeness to describe the rule sets $\mathcal{R}$, $\mathcal{T}$ and $\mathcal{E}$, which are subsets of $\mathcal{L}$.

A *function generator* $[\![.]\!]$ interprets a given rule set expressed in language $\mathcal{L}$ and outputs a function that maps elements of the universe $\mathcal{U}$ to real numbers in $[0, 1]$. It is used to generate functions that encode the rule sets $\mathcal{R}$ and $\mathcal{E}$, e.g. for all $c \in \mathcal{U}$: $[\![\mathcal{R}]\!](c) \in [0, 1]$. The function generator $[\![.]\!]$ is needed especially for transformational creativity, where the rule sets may change during the system's run and their interpretations need to be assessed again.

The *rule set* $\mathcal{R} \subset \mathcal{L}$ defines what kind of concepts are accepted as *valid* in terms of belonging to a certain class of objects such as mathematical theorems, buildings in a specific architecture style or songs in a specific genre. $\mathcal{R}$ can be used to define the conceptual space $\mathcal{C}$ using a validity threshold $k \in [0, 1]$. Validity threshold k is used to binarise if a given concept is valid enough to be accepted into the conceptual space, $\mathcal{C} = \{c \in \mathcal{U} \mid [\![\mathcal{R}]\!](c) \geq k\}$.

The *rule set* $\mathcal{E} \subset \mathcal{L}$ defines the *evaluation* function for the system, that is, the function generated by $[\![.]\!]$ through interpreting $\mathcal{E}$ allows to evaluate any concept $c \in \mathcal{U}$ as $[\![\mathcal{E}]\!](c) \in [0, 1]$. We define the set of *valued concepts* in the universe using a value threshold $l \in [0, 1]$: $\{c \in \mathcal{U} \mid [\![\mathcal{E}]\!](c) \geq l\}$.

The system's traversal of the conceptual space rests on a second *function generator* $\langle\langle ., ., . \rangle\rangle$. It takes into account the *traversal rule set* $\mathcal{T} \subset \mathcal{L}$, specifying how the system moves from concepts (or sequences of concepts) to other concepts (or sequences). Since traversal can also be informed by $\mathcal{E}$ and $\mathcal{R}$, the generator interprets all three rule sets, $\mathcal{T}$, $\mathcal{R}$ and $\mathcal{E}$ into a function which maps a sequence of input concepts, c_{in}, into a sequence of

$\mathcal{U}$	the universe containing all possible concepts
$\mathcal{L}$	a language in which to express concepts and rules, in a broad sense of the universe, $\mathcal{L} \subset \mathcal{U}$
$[\![.]\!]$	a function generator which maps a subset of $\mathcal{L}$ to a function which associates elements of $\mathcal{U}$ with a real number in $[0, 1]$.
$\langle\langle ., ., . \rangle\rangle$	a function generator mapping three subsets of $\mathcal{L}$ to a function that generates a new sequence of concepts of $\mathcal{U}$ from an existing one.
$\mathcal{R} \subset \mathcal{L}$	rules defining valid concepts
$\mathcal{T} \subset \mathcal{L}$	rules defining traversal in the concept space
$\mathcal{E} \subset \mathcal{L}$	rules defining evaluation of concepts

Table 2.1: Description of the elements in the Creative Systems Framework (CSF) [116].

output concepts, c_{out}:

$$c_{\text{out}} = \langle\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\rangle(c_{\text{in}}). \tag{2.2}$$

Similarly to Wiggins [52], we denote the set of all sequences of concepts reachable from given input concept sequence c using at most n recursive steps by

$$T^n(c) = \bigcup_{j=0}^{n} \langle\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\rangle^j(c). \tag{2.3}$$

With these elements, Wiggins defines characteristics for some interesting phenomena which may occur in creative systems on the object level, i.e. in exploratory creativity. Particularly, Wiggins [115, 116] introduces *uninspiration* and *aberration*.

Uninspiration Uninspiration occurs when a system fails to find valuable concepts, and it is defined with respect to the rule set $\mathcal{E}$. Wiggins distinguishes between three different types of uninspiration.

Hopeless uninspiration In hopeless uninspiration there are no valued concepts in the universe: $\{c \in \mathcal{U} \mid [\![\mathcal{E}]\!](c) \geq l\} = \emptyset$.

Conceptual uninspiration In conceptual uninspiration there are no valued concepts in the conceptual space: $\{c \in \mathcal{C} \mid [\![\mathcal{E}]\!](c) \geq l\} = \emptyset$.

Generative uninspiration In generative uninspiration the system cannot reach any valued concepts. From Equation 2.3 and given the empty concept $\top$ as the starting point: $\{c \in T^\infty(\{\top\}) \mid [\![\mathcal{E}]\!](c) \geq l\} = \emptyset$.

Aberration Aberration is defined based on how the valued concepts are related to the conceptual space and the system's ability to reach these concepts. Wiggins describes three types of aberration which all rely on the set of concepts which can be reached, but are not in the (current) conceptual space $\mathcal{C}$: $\mathcal{B} = T^{\infty}(\{\top\}) \setminus \mathcal{C}$.

Pointless aberration In pointless aberration none of the reachable concepts outside the conceptual space are valued: $\{c \in \mathcal{B} \mid [\![\mathcal{E}]\!](c) \geq l\} = \emptyset$.

Productive aberration In productive aberration a subset of the concepts in $\mathcal{B}$ are valued: $\{c \in \mathcal{B} \mid [\![\mathcal{E}]\!](c) \geq l\} \subset \mathcal{B}$.

Perfect aberration In perfect aberration all of the concepts in $\mathcal{B}$ are valued: $\{c \in \mathcal{B} \mid [\![\mathcal{E}]\!](c) \geq l\} = \mathcal{B}$.

Both uninspiration and aberration are interesting for a creative system as they define some detrimental modes from which the system may want to escape. To escape uninspiration the system may want to either change its evaluation standards (hopeless uninspiration), conceptual space (conceptual uninspiration), or its traversal (generative uninspiration). In productive and perfect aberration, the system may want to adjust its conceptual space to better accommodate aberrative concepts. In all of these cases, transformational creativity is required.

2.3.2 Transformational Creativity

Boden [8] defines transformational creativity as changing the conceptual space. In CSF [115, 116], transformational creativity is modelled as search on a meta-level, where object-level rules defining valid concepts $\mathcal{R}$ and traversal rules $\mathcal{T}$ can be adjusted. Generally, evaluation rules $\mathcal{E}$ may also be adjusted.

In CSF, the search in the meta-level consists of a similar septuple as the object level search defined in Equation 2.1:

$$\langle \mathcal{L}, \mathcal{L}_{\mathcal{L}}, [\![.]\!], \langle\!\langle ., ., . \rangle\!\rangle, \mathcal{R}_{\mathcal{L}}, \mathcal{T}_{\mathcal{L}}, \mathcal{E}_{\mathcal{L}} \rangle, \tag{2.4}$$

where $\mathcal{L}_{\mathcal{L}}$ is the meta-language and $\mathcal{R}_{\mathcal{L}}$, $\mathcal{T}_{\mathcal{L}}$, and $\mathcal{E}_{\mathcal{L}}$ are the meta-level rule sets for traversal. Without going into specifics of the formulation it is sufficient to say that this meta-level operates on object-level systems which are composed using language $\mathcal{L}$. $\mathcal{R}_{\mathcal{L}}$ defines what it means to be a creative system, $\mathcal{E}_{\mathcal{L}}$ defines how the *transformations* are evaluated, and $\mathcal{T}_{\mathcal{L}}$ defines how the search space of creative systems is traversed.

Using this two-layer structure between the object-level and meta-level systems, we can observe and investigate how the interplay between them may give rise to transformational creativity. For example, Grace and Maher [41] argue how specific curiosity on the object level, i.e. the habit of the producer to form an intention to explore some subspaces of the conceptual space when it encounters a surprising artefact, may facilitate transformational creativity.

Although CSF does not elaborate how exactly the information flows between the object-level and meta-level systems, we can assume that a creative system has some means to observe its state on the object level, e.g. with relation to aberrative behaviour, which then induces the meta-level to change the object-level system in some way. Of course, the more informed the system is of its current state and history, and the more powerful its meta-level is in its ability to come up with different kinds of object-level systems, the more informed and interesting transformations may take place.

The CSF provides a formalism for discussing and analysing exploratory and transformational creativity using a two-layer structure. However, it does not state how these are implemented in a particular system. Next, we move on to consider software architectures and metacreativity, which provide more tangible means to build (exploratory) creative systems and also enable transformational creativity.

2.4 Architectures of Creative Systems

Software architectures are one of the interests in software engineering. The term *software architecture* refers to the set of components included in the system and their relationships, i.e. the general design of the system. Descriptions of software architectures provide views on a system's operation where unnecessary implementation details have been excluded [90] and assist in grouping software with similar (architectural or design) *patterns* together [114]. Architectural patterns aid in understanding possible bottlenecks and other interesting factors in the design of the system, as the utilised patterns may cause characteristic operational fingerprints, e.g. in terms of maintainability [114]. Typical for architectural design is the *separation of concerns*, which roughly means that the system should be built modularly from components where each component has a well-defined responsibility and operational scope.

In creative systems, software architectures have not been subject to wide deliberate studies. As a subfield of AI, CC and the development of creative systems tends to focus more on the methods, which then determine the general architectural patterns used in individual systems. The design of the system may become tangled when the development is focused on the end products, and CC suffers from applied research not describing their system designs with shared vocabulary [39].

Given that some of the creativity theories explain creativity as a complex system of interacting and interrelated factors [30, 36, 57], it is striking that architectural aspects of creative systems have not been more readily addressed in CC. Although research using the cognitive approach has tied the general structures or layouts of creative processes to theories of consciousness [117, 119], not all CC systems *need* to derive their architectures from how the human mind (or body) works. The layouts of many AI methods do not bear any similarity to the (current knowledge of the) layouts and processes in the human mind, nor do they need to. Moreover, some researchers have put forward an idea of treating the source code as the first order object of interest in CC [24]. However, the proposition does not explicitly consider the creative system's architecture, rather it emphasises the relevance of studying creativity of the systems where the code is the (partial) end product generated by the creative system itself. That is, the end product is also a process that can be run by executing the code.

As CC research is free to use any layout or architecture it desires, a more concentrated effort on studying how architectural concerns affect the creative systems, and if there are any special concerns that need to be addressed when developing them, could prove to be beneficial. To put it more boldly, one could even say that architecture is the fifth perspective to computational creativity alongside the four Ps (Product, Producer, Process and Press). The four Ps have been derived from humans where architectural variations are not possible on the same scale as in computational systems, and thus do not need as close consideration.

The only paper (besides Paper I) that directly considers general architectural, or design, aspects of creative systems is written by Ventura [111]. Ventura outlines a structural design for creative systems which concentrates on how data flows within the system. The design distinguishes the following elements: domain knowledge, aesthetic, conceptualisation (genotype), genotypic evaluation, phenotypic evaluation and artefact (phenotype). The naming of the elements follows a similar convention as the terminology in genetic algorithms: the genotypic representation is the internal representation of the idea or concept, and the phenotypic representation is the (materialised) artefact that can be considered as an output of the system to the domain. Following the terminology in the paper, an artefact can be obtained from the genotypic representation using *translation*.

Ventura [111] sketches how the above-mentioned elements are connected to each other. Their general assembly to a single creative system bears similarity to the object level of CSF as there are no specific elements or components which would take care of the meta-level behaviour. Furthermore, Ventura introduces a procedure that can be followed by designers or researchers to build a CC system from scratch. The process starts by choosing the domain in which artefacts are produced, which advocates that the decisions made in the following steps are influenced by the chosen domain, rendering the process partially domain-specific.

2.5 Self-adaptivity, Self-awareness and Metacreativity

In this section, we cover our contributions to metacreative systems and their architectures [Paper I]. The contributions draw on previous research in software engineering, in particular self-adaptive and self-aware systems and their architectures [see, e.g. 55, 56, 58, 61, 95]. We first briefly introduce these concepts as they are defined in software engineering, and then proceed to our contributions.

Self-adaptive systems Self-adaptive systems are systems that can monitor and control their own behaviour. Often, self-adaptive architectures have a two-layer structure consisting of a *base system* and a *manager* [38, 55, 95]. The base system has a predefined higher-level goal, like producing creative end products or serving web pages to clients, and the manager component monitors how this goal is reached and provides adjustments to the base system if it observes that the base system's operation is defective or it could be improved in some ways. The adjustments may be as simple as changing to a different fixed configuration when the system observes a predefined context pattern, or they may require complex reasoning and recompiling parts of the system. The manager's observation and control processes can be realised, e.g. utilising a widely used Monitor-Analyse-Plan-Execute loop, where each of the elements may have access to a shared Knowledge base (MAPE-K loop) [55].

We have argued elsewhere that self-adaptation in software systems is in many cases inherently creative, and have separated three different creative adaptation types based on the way they are novel [67].

Self-aware systems Self-aware systems are systems that learn models of themselves on an ongoing basis and use those models in their reasoning enabling them to act based on their knowledge and reasoning [56, p. 5]. One of the main reasons for a system to learn the models is to affect its base system's behaviour during its execution, but also other reasons exists, e.g. adjusting what is reported to the maintainer of the system. That is, self-awareness can provide dynamic monitoring and control capabilities to a self-adaptive system, but not all self-adaptive systems are self-aware, as self-adaptivity may be realised using hard-coded rules.

The usage of the terms self-adaptive and self-aware systems is not completely fixed, however, and different researchers originating from different backgrounds may use them with slightly different connotations. A similar difference lies between the engineering and cognitive approaches in CC research when defining objectives of self-adaptation. The engineering approach aims for robust adaptivity methods, while research following the cognitive approach can be drawn to replicate adaptivity processes as they are understood in humans. Naturally, mixed approaches are common.

Self-adaptivity and self-awareness are closely related to *metacreativity* as it is defined in Paper I:

Definition 2. *A metacreative system is able to reflect and control its own creative behaviour.*

We combine reflection and control capabilities from Definition 2 as *self-awareness* in Paper I, which is not exactly the same definition as in Kounev et al. [56, p. 5] cited above. As our interest is in adjusting the creative behaviour by monitoring it, we only consider systems where the control capabilities are geared towards the creative behaviour itself. Moreover, we do not require ongoing learning capabilities from all metacreative systems but differentiate between systems with fixed connections from reflection to control (coined *exogenous connections* in Paper I) and systems that are able to learn these connections (coined *endogenous connections* in Paper I) during their execution. In many cases, creative applications and phenomena are unpredictable or even chaotic, which places high demands on learning the connections, but at the same time emphasises their need as the system's developer can not anticipate the situations the system encounters during its execution. As such, metacreative systems are by definition self-adaptive, with a strong emphasis on self-awareness as defined by Kounev et al..

We distinguish between different self-awareness aspects based on their awareness target (Figure 2.1). The awareness aspects are artefact-awareness, generator-awareness, goal-awareness, interaction-awareness, time-awareness and meta-self-awareness. For example, an artefact-aware system is able to control and reflect on the end products it creates. The

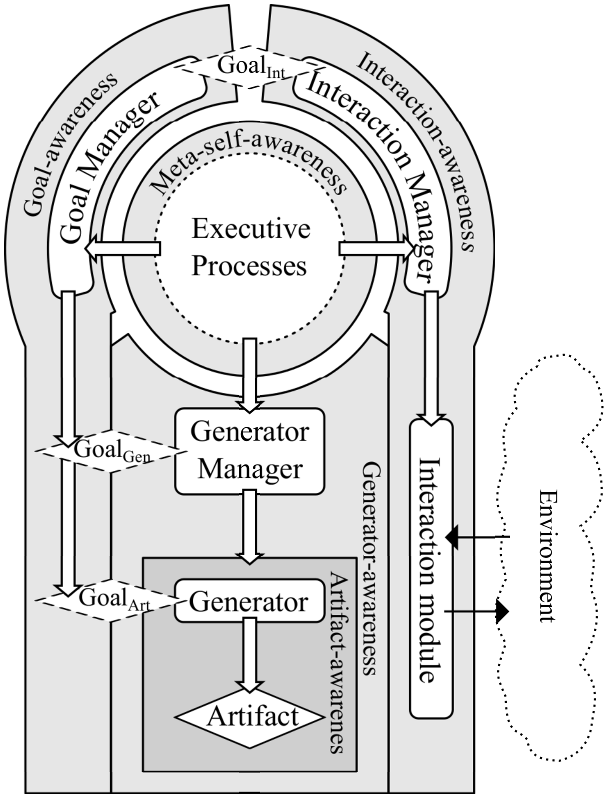

Figure 2.1: Different self-awarenesses described in Paper I and their stereotypical connections. White arrows point to the targets of reflection and control and the blunt end marks the component responsible for that behaviour. Time-awareness is omitted and only the most common relationships are shown.

distinguished awarenesses are inspired by Lewis et al. [61], but we modify the awareness set to suit creative systems and consider their impact, particularly in the context of creative systems.

By combining different awarenesses, we can provide different characterisations of creative systems from the Producer perspective, where metacreativity is the unifying characteristic. These characterisations can in turn be used in the analysis of creative systems. For example, a goal-aware and artefact-aware system which changes its goals based on the information it gains from reflecting on the artefacts can be said to be autonomously creative as defined by Jennings [49].

The stereotypical layout of how these awarenesses are connected to each other is shown in Figure 2.1. In the figure, we can see that the awarenesses have diverse relationships, of which only the most common are shown. Particularly, time-awareness could be connected to any component but is omitted from the figure.

Even though Figure 2.1 may appear complex, these elements can, in general, be fitted to the two-layer structure of self-adaptive systems, i.e. a base system and a manager with MAPE-K loop. The base system would consist of a generator that produces artefacts and a possible interaction module (and other elements such as a knowledge base, etc., as discussed by Ventura [111]), and all the managers and executive processes would be in the manager component. On the other hand, the executive processes inside meta-self-awareness could also be thought of as a third layer that combines the information from the individual manager components into a unified view of the creative system's status.

Chapter 3

Creative Agents

Creative agents are focal to Computational Creativity (CC). For example, a chatbot, a robot keyboardist developed for jazz improvisation with human musicians, a computational painter choosing which strokes it performs on a canvas and a robot dancer can all be modelled as agents.

In this chapter, we cover fundamentals of creative agents and our contributions to creative action selection [Paper II] which extends the Creative Systems Framework (CSF) [115, 116] (see Section 2.3). The chapter also provides background for Chapter 4 where we focus on multiple creative agents operating in the same environment. Similar to Chapter 2, we handle creative agents mostly in an abstract manner without particular agent implementations in mind. The most eminent of the four Ps in the chapter is the Process perspective as *how* the agent decides *what* to do next is pivotal for the agency.

We start by formulating our definition of a creative agent, after which we briefly discuss general creative processes. Then, we cover the fundamentals of the sequential action-selection problem in AI agents, i.e. Markov Decision Processes (MDPs), and outline how this problem may be solved using Reinforcement Learning (RL). Lastly, we present our contributions to action selection for creative agents which synthesises concepts from the action selection in AI agents into the CSF.

3.1 Definition of a Creative Agent

Similar to creative systems and their definition (see Section 2.1 and Definition 1), there is no single agreed definition of a creative agent. There has not even been a concentrated effort towards an integrated definition of a creative agent in CC. This is not particularly surprising, as what is a single (non-creative) agent (in AI) has not been answered unambiguously either [100, p. xvii]. In philosophy, sociology and ethics *agency* is defined as the capability of an entity to act in its environment. However, the position and quality of the boundary separating the agent and its environment are under philosophical debate [48].

We adopt a pragmatic approach where an agent is separated from its environment with a clear distinction of which elements constitute an agent and which are part of the environment (or other agents). Due to this separation, the Press perspective is always at least implicitly present when discussing creative agents. However, as the Press perspective deals with who gets to evaluate creativity and why, we leave the majority of the discussion about if for Chapter 4 where we focus particularly on social creativity.

The agent can attempt to manipulate or have an effect on the environment by performing *actions*. The agent is free to choose which action it performs at any given time, but the actions available to the agent may be restricted because of the current *situation* or *context*. By situation, we refer to the combined current state of the environment and the agent, and by context, we refer to the state of the environment.

For the needs of this thesis, we use the definition of a creative agent similar to Paper II:

Definition 3. *A creative agent is a utility-based agent with a goal to produce creative, i.e. novel and valuable, concepts and/or artefacts.*

Definition 3 combines the concept of *intelligent agents* by Russel and Norvig [93] with the standard definition of creativity (see Section 1.2). It considers creative agents to be utility-based, which means that a creative agent has some form of heuristics or other means to measure how close it is to the desired solution, not just a binary understanding if it has reached a creative end product or not. To simplify, the utility can be understood to be a linear combination of novelty and value, although they can also be measured separately and an agent's reasoning can be informed of both measurements. Furthermore, arguments have been raised to question if the novelty of the stimuli, i.e. perceived artefact, should be measured directly, i.e. the difference of the artefact to the previously encountered artefacts, or should the agent's *response* (or arousal) to the novel stimuli be measured instead [6]. The latter point of view is adopted often when researching curious behaviour in creative agents [see, e.g. 98].

Our definition of creative agents references both concepts and artefacts, aiming to explicate the difference between the agent searching for interesting inner ideas, i.e. concepts, and materialising these inner ideas as artefacts into the environment. Crucially, a concept may be materialised as multiple artefacts and one artefact may be observed to express multiple concepts. This concept-artefact distinction is present both in theories of human creative processes [74, 88, 94] and in previous Computational Creativity studies [cf. 41, 111].

3.2 Creative Processes

Creative processes become central when changing focus from creative systems to creative agents as the agency point of view embraces creative systems as actors in their environment. This emphasises questions like "What are the properties of the actions the agent chooses during its operation?" and "Why does the agent choose particular actions?". Moreover, the issue about creativity's domain generality [71, 86] becomes particularly apparent when dealing with creative computational processes, as many domains have their particular methodologies for producing artefacts. Similar evidence is also found in humans and specifics of the artistic creative process [10].

Below, we discuss a few selected theories of creative processes previously discussed both in the context of human creativity and Computational Creativity. The aim is to give a rough overview of typical phase successions that occur during creative behaviour and how creative behaviour has been previously modelled in CC. For computational processes, we provide overviews of Markov Decision Processes and Reinforcement Learning (see Section 3.3) as they are directly relevant for this thesis. In particular, we do not cover all processes or process types used in artificial, creative agents (or systems) as they are in many cases implementation specific. For example, we do not cover other machine learning processes which could be deemed creative, e.g. learning a novel weight distribution for a generative adversarial network's (GAN) generator which is able to produce novel and realistic images in predefined categories [11, 40].

In humans, the creative process is a sequence of thoughts and actions [72]. This succession is often modelled as what could be called a *stage theory*. In a stage theory, the creative process is divided into successive stages, where some (subset of) stages may be repeated iteratively until an acceptable solution is found.

A typical example of a stage theory is the first theory of the creative process proposed by Wallas, later studied by Sadler-Smith [94], where the creative process is divided into four stages: *preparation*, *incubation*, *inspiration* or *insight*, and *verification*. In preparation, the individual defines the problem; in incubation, the individual withdraws from the problem and the process continues subconsciously; in insight, the individual experiences a moment of 'Heureka' and a solution to the problem seems imminent; and in verification, the actual solution is formulated and tested. Sadler-Smith [94] argues that Wallas included a fifth stage, intimation (fringe consciousness), between the incubation (non-consciousness) and inspiration (consciousness) stages. Others have proposed that there are also other stages, e.g. between preparation and incubation lies a stage of conscious concentration [10, 74] where the problem is actively worked on.

After Wallas, multiple stage theories with complex relationships between each other have been proposed. Botella et al. [10] report 20 different stage theories, identify and localise 16 different successive stages combined over all the theories (such as idea generation, realisation, etc.), and map how the stages of each theory are related to the stages in other theories. In their analysis, the same stage in the combined stage succession over all theories may have different names and meanings in different theories.

Botella et al. [10] characterise stage theories as describing the creative process on a *macro level*, i.e. they define what happens overall during a creative process. On a *micro level*, underlying mechanisms of the creative process can be discussed, such as divergent and convergent thinking.

In CC, the concentrated study of the special properties of creative processes has not raised wide-reaching interest, and only a few scholars have studied more general theoretical aspects of creative computational processes which go beyond implementations in individual systems.

Boden's [8] three creativity types (see Section 2.2.1) characterise different creative process types, but it does not answer why the system would exhibit, e.g. transformational creativity, or what exactly the actions the system takes are, e.g. when exploring the conceptual space. Moreover, the FACE model [19, 81] specifies different generative act *types* (see Section 2.2.5), e.g., generating a concept and generating a procedure that can generate new aesthetic measures. However, because of its scope, the FACE model handles generating an end product as single action and not as a succession of multiple actions which could be further analysed.

The CSF [115, 116] (see Section 2.3), on the other hand, can describe some unwanted modes of the object-level system, e.g. aberration, which the meta-level system can observe as a trigger for a transformation, but it also treats the whole exploration process (on both the object level and meta-level) as a black box. For example on the object level, traversal between concepts is realised by the interpreted traversal function (Equation 2.2) which inner operation is left (intentionally) undefined. In particular, the CSF does not explicate in detail how the set of traversal rules $\mathcal{T}$ is used to influence to which output concept c_{out} the system moves next.

Schmidhuber [99] has proposed a formal theory of a human developmental process. The theory argues that creative behaviour can be explained by a simple mechanism that utilises Reinforcement Learning (RL) to maximise *the fun* for the discovery or creation of *novel patterns*. Sensory perceptions, or data in general, contain a pattern when they are compressible. The main interest is in *learning* the patterns from sensory perceptions. The learning can be translated into an *intrinsic reward*, i.e. to a reward which is personal to the agent and is not given from the environment, aiming to capture how much better the agent's world model can explain its observation history in terms of some performance criteria after the learning takes place. This intrinsic reward is understood as internal joy, or fun the agent is having, and relates to the motivation of the agent; too novel and too ordinary patterns do not rouse internal joy as the agent either does not learn or is not able to learn anything from them. The performance criteria can be defined, e.g. as the world model's compression performance on the observation history. The idea is that using this kind of mechanism, the RL controller is continuously driven to produce data (sensory perceptions) which is surprising, and by learning the patterns in that data the agent improves its world model and may acquire more complex behaviours.

Overall, the implementation specificity of creative systems and agents makes it challenging to provide a unified framework of the creative process in computational systems. Moreover, the implemented systems typically only focus on one or two stages of the human creative process, namely *idea generation* (generating concepts) or *realisation of the ideas* (generating artefacts) and *verification* (evaluating ideas or artefacts using computational means). In this sense, the majority of the proposed CC systems and agents can be understood as being implementations of micro-level processes. The designer of the system guides the macro-level process by developing the computational system, e.g. for idea generation, and validates the results, e.g. by using some of the analysis tools introduced in Chapter 2.

Next, we provide an overview of two widely used means to model and solve how intelligent computational agents can select useful actions, Markov Decision Processes (MDPs) and Reinforcement Learning (RL). These theories and their implementations provide background for our contributions to modelling the action selection process in creative agents (see Section 3.4).

3.3 Action Selection Problem and Solving It

The action selection problem for intelligent computational agents is a well-studied topic in AI [see, e.g. 87, 106]. It is often modelled as a Markov

Decision Process (MDP) or Partially Observable Markov Decision Process (POMDP) [87]. Below, we cover the fundamentals of MDPs and discuss briefly how Reinforcement Learning can be used to approximate solutions for them. We do not cover POMDPs or theoretical details of RL, as they are not central to this thesis.

Our extension of the CSF for the action selection process of creative agents [Paper II], discussed in Section 3.4, borrows concepts from MDPs and their solutions. The MDPs are also implicitly present in our simulations of creative agent cooperation in Paper IV and Paper V: the agents utilise simple RL methods which are based on MDPs during simulation development, even though MDPs are not mentioned in the papers.

3.3.1 Markov Decision Processes

A Markov Decision Process (MDP) defines a *sequential decision making problem* where an agent interacts with its environment over time. On each time step t the agent may receive a reward (from the environment), which is potentially affected by the agent's *actions*. The environment is distinguished from the agent by everything the agent cannot change arbitrarily, and as such, it fits our pragmatic view of the agent and the environment having a clear boundary.

A (finite horizon) Markov Decision Process is defined as a quadruple

$$\langle \mathcal{S}, \mathcal{A}, p, \rho \rangle, \tag{3.1}$$

where $\mathcal{S}$ is the set of all possible environment states and s_t, where $s_t \in \mathcal{S}$, is the state of the environment on time step t; $\mathcal{A}$ is the set of actions available to the agent and a_t, where $a_t \in \mathcal{A}$, is the action the agent performs on time step t; $p : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ is the *environment dynamics*, which defines how the actions the agent performs potentially affect the environment state, given by a conditional probability distribution $p(s_{t+1} \mid a_t, s_t)$; and $\rho : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$ is the *reward dynamics*, which defines how the movement between environment states using particular actions are rewarded, given by $\rho(s_{t+1}, a_t, s_t) = r_{t+1}$, where r_{t+1} is the reward on time step $t + 1$. Markov assumption implies that all the needed information of the past interaction between the agent and the environment is encoded in the current state s_t.

The decision-making behaviour of the agent, or an attempt to solve a Markov Decision Process, is defined by a *policy* π, $\pi : \mathcal{S} \times \mathcal{A} \to [0, 1]$, which is given by a conditional probability distribution $\pi(a_t \mid s_t)$. That is, for each state $s \in \mathcal{S}$ the agent may have a different emphasis for the action it should take.

A *solution* to a Markov Decision Process is an *optimal policy* π^* which maximises the expected future reward $V^\pi(s)$ when following a policy π and starting from state s:

$$V^\pi(s) = \mathbb{E}\left[\sum_{k=0}^{\infty} \gamma^k \rho(s_{t+k+1}, a_{t+k}, s_{t+k}) \middle| s_t = s\right], \tag{3.2}$$

where $0 < \gamma \leq 1$ is a discount factor that defines how near-future rewards should be weighted in comparison to rewards in the more distant future. Crucially, a Markov Decision Process may have multiple optimal policies.

Markov Decision Processes (and POMDPs where the agent may not observe the full environment state on each time step) offer a flexible framework for defining different action selection problems for agents. However, they do not state how the agent should come up with its policy π. Multiple different methods can be used to devise an agent's policy starting from a human designer handcrafting it, however, many of the methods fail when the problem space becomes too complex.

3.3.2 Reinforcement Learning

Reinforcement Learning (RL) offers a practical way to solve action selection problems by *learning the agent's policy from its experience* [106].

In Reinforcement Learning, an agent interacts with its environment in episodes and adjusts its policy either after each episode or simultaneously during the interaction. Some of the agent's actions during an episode may include *exploration*. That is, the agent does not need to follow its policy on every action. It may explore the state-action space in order to come up with better policies.

The learning can be *off-policy* or *on-policy*. In off-policy, the (approximated) value of the optimal policy π^* is learned independently of the actions taken by the agent (e.g. Q-learning [113], which is utilised in Paper IV and Paper V). In on-policy, the agent learns the value of the current policy it follows (e.g. State–Action–Reward–State–Action (SARSA) [91]).

Reinforcement Learning methods can be divided into *model-free* and *model-based* methods [106]. In the model-free methods, the agent directly learns either state values for their policies, V^π, as shown in Equation 3.2 or, more often, state-action values, $Q : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$, where $Q(s, a)$ is the value of an action a taken in a state s. An example of a model-free method is Q-learning. In the model-based methods, the agent is either given or learns a "world model", where the model aims to capture relevant aspects of environment dynamics possibly combined with reward dynamics. However, combining the learning of environment dynamics and reward dynamics makes the model challenging to use in different tasks with different rewards. For example, Schmidhuber's [99] proposition of a human developmental process is a model-based RL method as the agent specifically learns a model of the world.

3.4 Creative Action Selection Framework

While Markov Decision Processes and Reinforcement Learning are general frameworks for action selection and learning from experience, they do not explicitly take into consideration specifics of creativity. Moreover, even though RL processes can be deemed creative in many cases [see, e.g. 15, 112], not all creative systems which can be modelled as agents utilise them. Many creative systems use heuristics or are inspired by other frameworks, and can still be understood to interact with an environment and choose actions in a sequence. That is, they fulfil the basic requirements for our definition of creative agents.

In Paper II, we aim to provide the first steps towards a formal framework for creative action selection in the context of analysis of creative systems or agents. The framework, dubbed the Creative Action Selection Framework (CASF), extends the CSF (see Section 2.3) and borrows, or synthesises, elements from MDPs into it.

In particular, the CASF dissects CSF's traversal function (Equation 2.2) into constituents that are relevant for creative agents, and provides some stopping criteria for them. Moreover, we make time and concept-artefact separation explicit in the framework, and associate point $c \in \mathcal{U}$ in the CSF, i.e. a (partial) concept or an artefact, to a state in MDPs. Lastly, we distinguish different subsets of traversal rules $\mathcal{T}$ for the actions of the agent, its world model (however, not all agents need to have a world model in the CASF) and the policy it follows. We consider as action any choice the agent can make during its operation.

The action selection loop of a creative agent according to the CASF is illustrated in Figure 3.1. We scrutinise how each of the elements *observe*, *assess*, *filter*, *predict*, *choose* and *execute* are part of a creative agent's action selection and provide a general mathematical formalism for each function.

The four functions constituting the action selection (choose, assess, predict and filter) are separated for conceptual clarity. In many cases, an agent may not have all of these functionalities, or they are entangled within different processes of the agent. Moreover, their order in specific agent implementations may differ.

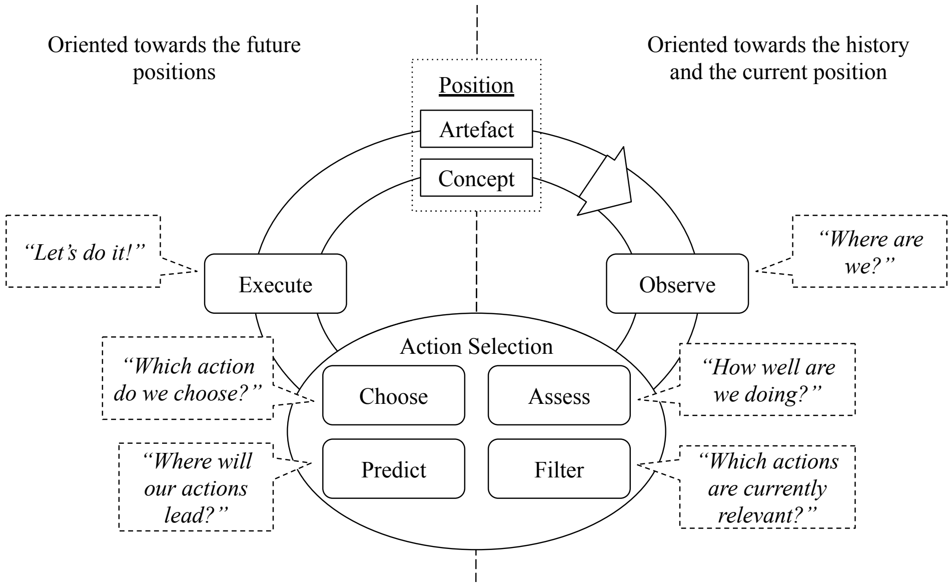

Figure 3.1: Creative action selection loop as depicted in Paper II.

Using this conceptualisation of the action selection loop of a creative agent, the CASF aims to provide a shared vocabulary that can be used to analyse processes in diverse creative agents, utilising RL or other methods. Especially, we can distinguish different properties of the action capabilities. These can be simple statistical variables such as what the number of available actions in any given situation is, or more general characteristics such as whether the agent is able to learn new actions during run time (Producer perspective). Statistical variables can also be observed over time, e.g. if the agent's action selection after filtering suitable actions in each situation grows and shrinks in a controlled manner over time, then one could hypothesise that the agent's creative process has similarities to convergent and divergent thinking. A higher-level analysis comes from the state-action sequences the agent performs (or is capable of performing) and what their properties are. For example, we can observe if the state-action sequence always has a certain form or if the form differs between the creation of each end product (Process perspective).

Overall, the CASF extends the analysis capabilities of the CSF for creative systems which can be modelled as agents. It models the creative process as a sequence of state-action pairs where the action selection process of the agent is informed of the end products it is aiming to create, providing detailed analysis capabilities for the Process perspective. In comparison to Markov Decision Processes (and Reinforcement Learning), the CASF deliberately distinguishes between concept and artefact spaces and other elements relevant for creativity, e.g. rules for validity $\mathcal{R}$, making it easier to discuss the specifics of the creative process.

In Paper II, the CASF is described mostly from the engineering approach, but it can also be used in studies with the cognitive approach as well if properly operationalised. Moreover, the generality of the states and actions makes it possible to describe creative processes both on a micro and macro level. Some actions may require fine motor skills to move a robot's brush on a canvas while higher-level actions may mark the movement between the overall stages of a creative process, e.g. from idea generation to realisation. For example, (some of) the action types specified in the FACE model [19, 81] could be either single actions or composed of multiple actions in the CASF.

Chapter 4

Social Creativity

In this chapter, we cover aspects of social creativity relevant for this thesis, and our contributions to both theoretical analysis of creative agent societies [Paper III] and empirical collaboration simulations in the abstract art domain [Papers IV and V].

Social aspects of creativity accounted by the Press perspective are an important part of creative phenomena: all creative agents operate in some environment, which typically contains other (creative) actors or entities. When shifting the focus from single agents to multiple creative agents operating in the same environment, attention is drawn to the *interactions* between the agents. These interactions can result, e.g. in *collaboration* potentially allowing the agents to produce something together they could not create alone [109].

We concentrate on *multiple computational agents interacting with each other* and do not consider co-creative systems where a computational system interacts with a human user. We base our inspection of social creativity in computational entities on theories of social creativity in humans, namely the Systems View of Creativity [30], which is also called the Domain-Individual-Field-Interaction model (DIFI model) [35].

Our research approach to social creativity is Computational Social Creativity (CSC). Research on CSC is primarily based on Agent-based Models (ABMs) [34], which are described more as a mindset than a technology [9]. The main point of view is that in an ABM the whole system, i.e. society, is described from the perspective of its constituents, i.e. agents. The focus is on simulating the agent society and observing what kind of *emergence* arises from different agent properties and interactions. This is a stark contrast to Multi-agent Systems (MASs), where a typical goal is to make a set of agents operate as effectively as possible in a particular problem domain [100].

We begin by introducing the DIFI model [30, 35]. We continue by describing Computational Social Creativity and briefly cover previous work in its field. Then, we present our contributions to merging the DIFI model with the Creative Systems Framework [Paper III]. Lastly, we cover collaboration as a special case of social creativity and our contributions to empirical collaboration simulations [Papers IV and V].

4.1 The Systems View of Creativity (DIFI model)

The Systems View of Creativity introduced by Csikszentmihalyi [30], and later developed into the Domain-Individual-Field Interaction model (DIFI model) [35], is a prominent theory dealing with social aspects of human creativity. Its main argument is that *creativity is not isolated in any individual*, but instead, it is best understood in its entirety by investigating the dynamics between the *individuals* of the society, the experts of a given *field* and the *domain* of accumulated cultural artefacts, as depicted in Figure 4.1.

The DIFI model encompasses all of the four Ps, with an emphasis on the Press perspective, and thus it attempts to capture the high-level account of the dynamics of diverse creative phenomena within a society. The DIFI model consists of the following elements:

Individual An individual is a human or other entity, e.g. an artificial creative agent, which operates in some environment. The individual has its own characteristics, traits and skills (Producer perspective) and it follows its own creative processes (Process perspective on an individual level) to produce end products it deems creative by its own standards (Product perspective, in the context of P-creativity or small-c).

Field The field is a collection of social institutions relating to some class of artefacts, e.g. in contemporary art, in a scientific discipline such as CC, or in jazz music. The individuals, which we here call experts, belong to the field if they can affect what is preserved in its domain. These experts in the field may be (and often are) individuals who produce novelty in that same field, but they can also be, e.g. critics, teachers and collectors. The same individual may belong to multiple fields in multiple roles and the fields may overlap each other. In some cases, the field may be implicitly defined, and the members of the field may not recognise themselves as such.

The field acts as a filter for the artefacts which are produced by the individuals, and select which artefacts are accepted to the domain of cultural artefacts (Press perspective). In some fields this selection mechanism can

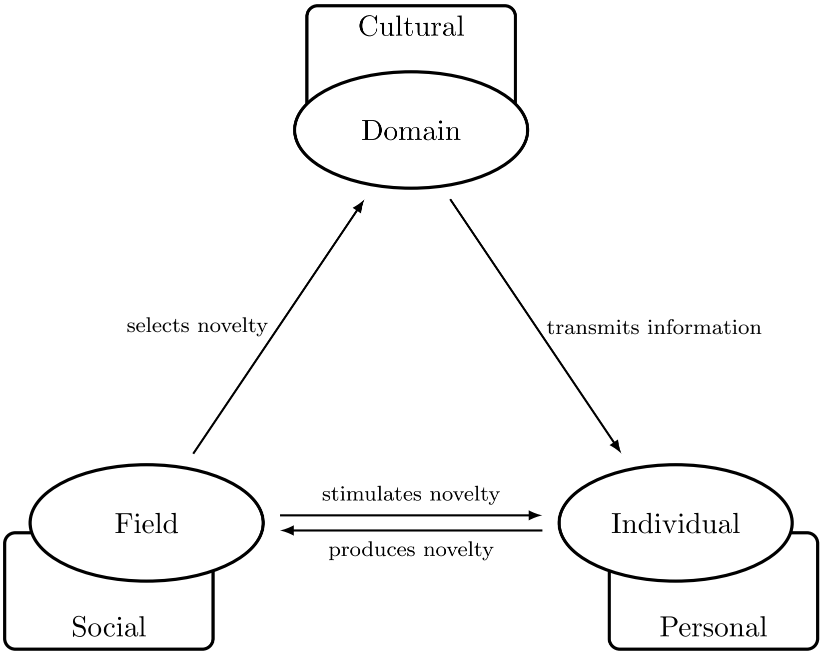

Figure 4.1: The relationships between different elements of the DIFI model, adapted from Csikszentmihalyi [30]. An *individual* produces novelty, which is assessed by the experts in the *field* to be accepted to the *domain* of cultural artefacts. The field stimulates the individual's generation of new end products while the domain transmits information about the culturally accepted artefacts to the individual.

be explicit, e.g. the peer review process for scientific publications follows specific standards, and in others implicit or fuzzy, e.g. street graffitis which are seen to hold exceptional value are less likely to be drawn over by other street artists (Process perspective on a societal level).

The field, in general, can act in multiple spatial or cultural abstraction levels from local galleries to culturally accepted seminal works. Thus, the field may deploy different selection mechanisms depending on the context of the novelty produced by the individual. For example, a local gallery curator may select artworks to the gallery based on its own assessment, whereas artworks that are seen as culturally transformative are usually selected by the field as a whole after a period of time from the creation of the artworks has passed and the field has continued to develop.

Domain The domain is a collection of cultural artefacts which have been selected by the field (Product perspective, in the context of H-creativity or Big-C). In a similar way as the selection of the field functions, the domain may have multiple abstraction levels from small local galleries to distinguished museums with vast collections or from local symposiums to leading scientific journals. Moreover, the importance of different cultural artefacts may be observed in varying ways in different locations, at different times, or by individuals with different cultural connections and personal bonds.

Interaction All the three elements above, individual, field and domain, are continuously interacting with each other. The domain transmits information to the individual as the individual observes its artefacts, and the individual brings about some variations based on its own observations and impressions of the domain artefacts, using its own creative processes and assessment of the artefacts. These variations are then proposed to the field which utilises its own processes to select the variations which are accepted into the domain.

The field or the social context can also act as a *generative force* for the individuals. For example, some individuals which would not otherwise produce creative artefacts may be drawn to do so when the social context is appealing [30].

4.2 Computational Social Creativity (CSC)

The studies of social creativity in CC can often be categorised in the general context of Computational Social Creativity (CSC) [96], which roughly refers to studying creativity with Agent-based Models. In its approach and methodology, CSC is close to fields such as artificial life and computational social sciences. In this thesis, Papers III–V all adopt CSC as their general context. Paper III aims to provide descriptive analysis tools for CSC research, and Papers IV–V study, among other things, emergence in a society of creative agents.

CSC has four properties which are either required (the first two) or desirable (the last two) [96]:

1. **Models must demonstrate a mechanism** CSC's approach to modelling is qualitative and the main focus is on the discovery of mechanisms. Similar to computational social sciences, the models should be generative in nature [26]. That is, the models are constructed using beliefs or knowledge of how individuals behave and

the simulations are then executed and the results investigated. Different individual traits, characteristics or interaction principles may produce different behaviour on a macro level, i.e. the society exhibits *distinct emergent qualities* based on its individuals.

2. **Models must be simple and reproducible** The main claims of CSC models should be reproducible by other researchers using different source codes than the original study used. Thus, using simple models allows the researchers to focus on the core aspects of the model, making it easy for others to reproduce the effects. However, Saunders and Bown [96] acknowledge that CSC studies highly complex behaviour, which may be challenging to represent or reproduce in simple models.

3. **Models should preferably generate new hypotheses** As CSC offers a platform to study hypothetical creative situations, contexts, individuals and interactions, it can be used to study creativity as it could be instead of as it is. This can promote new hypotheses considering creativity, which may be further investigated using CSC as the platform.

4. **A strong CSC model would actually be creative and achieve the goal of CC** Even though CSC focuses on simple individuals which are typically minimally intelligent, a *strong* CSC model would make the whole society appear creative on a macro level. This claim is not as far-fetched as it may sound, as *emergence* can be seen as creative behaviour in its own right: it can produce a novel behavioural pattern that is useful, especially for the individuals in the society.

Theories of social creativity, such as the DIFI model [30, 35], can be incorporated into CSC studies either during the modelling, e.g. how critics explicitly decide which artefacts are accepted into the domain, or during the interpretation and analysis of the results, e.g. what kind of implicit fields are formed during the simulations with certain interaction behaviours or different (distributions of) individual characteristics. However, the research setting has to be properly devised and the interpretation of the results has to be done carefully to separate causation from correlation.

4.3 Previous Research in CSC

Previous research that can be roughly categorised into the context of CSC is diverse with many different application domains and investigations of interesting mechanisms. Below, we present a few selected examples of CSC research to outline the field's general research interests. For a more thorough account of previous studies in social creativity, we refer the reader to Saunders and Bown [96].

Saunders and Gero [97] study societies of agents that produce genetic artworks. In their simulations, the agents are curious, i.e. they appreciate novelty. Each agent converts the novelty value into interestingness using the Wundt curve [6]. The agents prefer different levels of novelty, i.e. their interestingness peaks at different levels of novelty. Saunders and Gero observe the emergence of cliques where agents appreciating similar levels of novelty appreciate each other's artworks more often.

Sosa and Gero [102] study design creativity in the context of the DIFI model. They discover multiple interesting phenomena emerging during the agent society simulations. For example, they observe *dissenters*, agents that may trigger change cycles in society in a bottom-up style, and the emergence of *gatekeepers*, agents which have an emphasised role in the field's collective decision to accept artefacts into the domain.

Gabora and Tseng [37] study how self-regulation affects the overall creativity of society. In their simulations, the agents can invent new or imitate existing actions. The actions consist of moving different body parts and agents prefer symmetrical movements. They conclude that self-regulation, i.e. the individual's tendency to imitate existing actions instead of inventing new actions if it has not been able to invent fit actions in the past, benefits the whole society. The actions become more fit in the whole society when compared to a non-regulated society and the diversity of the actions is higher, especially during the beginning of the simulation.

We have previously investigated social decision making using novelty-seeking agents which produce spirographs in iterative simulations [68]. In particular, we studied how self-criticism and veto power affect the artefacts which are accepted into the domain as a result of all agents voting for the most novel artefact on each iteration. We observed that self-criticism is beneficial in agent societies where collective voting is employed as it drastically reduces the number of artefacts each agent needs to assess. Moreover, veto power and self-criticism both raise the average novelty of the artefacts accepted into the domain, while too aggressive use of veto power raises the efforts needed to produce the same number of artefacts for the domain.

Overall, the typical CSC approach uses (iterative) simulations where different properties of the agents or the simulation setups change. The results are either compared between different simulation setups or between different agents inside the same simulation run. As a rule, the simulation setups are run multiple times to get statistically relevant results. The different agent properties may be as simple as different labels of agents where agents are biased against agents with different labels, or they can consist of properties such as the resources each agent has, the number of social connections available to each agent, and the agent's ability to reason and learn. The simulation setups may vary by giving all the agents different properties or resources, changing the society from sparse social connections to dense social connections, or varying the number of agents in the simulation.

4.4 Formal Analysis of Creative Agent Societies

In this section, we cover our contributions to the formal analysis of creative agent societies as proposed in Paper III. The paper merges elements from the DIFI model (see Section 4.1), namely individuals, the field and the process of selecting artefacts into the domain, into the Creative Systems Framework (see Section 2.3). The resulting extension to the CSF, dubbed the Creative Systems Framework for Creative Agent Societies (CSF4CAS), is developed to be used as a formal analysis tool to research social creativity, e.g. in the context of Computational Social Creativity. Principally, the CSF4CAS is also compatible with Liu's dual generate-and-test model [69].

The framework defines a society $\mathbf{S}$ as a set of creative agents, $\mathbf{S} = \{A_1, A_2, \ldots, A_n\}$, where each agent A_i has its own *independent* object level septuple of the CSF (see Section 2.3 and Equation 2.1). As agents may change as time passes, we denote the CSF of agent A_i on time step t as

$$\langle \mathcal{U}, \mathcal{L}, [\![,]\!] \langle\!\langle ., ., . \rangle\!\rangle, \mathcal{R}_{A_i^t}, \mathcal{T}_{A_i^t}, \mathcal{E}_{A_i^t} \rangle,$$

where $\mathcal{R}_{A_i^t}$, $\mathcal{T}_{A_i^t}$ and $\mathcal{E}_{A_i^t}$ are agent's A_i rules for validity, traversal and evaluation on time step t, respectively. We denote by $R_{A_i^t}$ the set of artefacts considered valid by A_i on time step t, $E_{A_i^t}$ the set of artefacts considered valuable by A_i, and $T_{A_i^t}$ the set of artefacts reachable by A_i using any number of applications of the traversal function (see Equation 2.3). For simplicity, we omit defining the time step when it is not necessary.

The object-level independence means that while the agents may influence each other, only the agent itself controls its own creative behaviour defined by the septuple. In general, the agents conform to our definition of creative agents (Definition 3): in the CSF the utility measure can be defined as a linear combination of evaluation and validity functions $[\![\mathcal{E}_{A_i}]\!]$ and $[\![\mathcal{R}_{A_i}]\!]$, respectively. Although, basing it solely on $[\![\mathcal{E}_{A_i}]\!]$ would be a typical solution in CSC as $[\![\mathcal{R}_{A_i}]\!]$ defines the current conceptual space of *acceptable solutions* for agent A_i.

In Paper III we do not make an explicit separation between concepts and artefacts nor do we consider different components of the evaluation function, but the CSF4CAS could incorporate into itself the single-agent CSF elaborations of Paper II as well.

In the centre of the CSF4CAS is an abstract *societal aggregation function* Π which takes as input a set of agents $\mathbf{S}$ and computes from input *societal representations* for the set of artefacts valued by society $E_{\mathbf{S}}$, the set of artefacts seen as valid by society $R_{\mathbf{S}}$, and the artefacts reachable by society $T_{\mathbf{S}}$. That is, for society $\mathbf{S}$ on time step t:

$$\Pi(\mathbf{S}^t) = \{R_{\mathbf{S}^t}, T_{\mathbf{S}^t}, E_{\mathbf{S}^t}\}, \tag{4.1}$$

where $R_{\mathbf{S}^t}$ denotes what the set of agents $\mathbf{S}$ socially considers to be the conceptual space of acceptable solutions on time step t, i.e. it determines the *field* the set of agents agree on; $T_{\mathbf{S}^t}$ denotes the artefacts reachable by the set of agents, i.e. the exploratory capabilities of the whole society; and $E_{\mathbf{S}^t}$ denotes the end products the set of agents socially evaluate to be good enough, i.e. which artefacts pass the social creativity test of the field and can be included into the domain. That is, Π incorporates into itself *the social decision-making processes of the given field and its individuals*, which may be either explicit or implicit as discussed in Section 4.1. The implementation of Π is society specific, and may be one of the main interests of the research: what kind of societal phenomenon is observed with different societal aggregation function Π implementations?

The CSF4CAS obscures all the particular interactions between the agents from its analysis. This may seem counter-intuitive when compared to the general interests of CSC where the interactions are one of the main aspects of the study. However, CSC models should also demonstrate mechanisms (the first requirement of CSC, see Section 4.2) and CSF4CAS allows to analyse the effects different implemented interaction mechanisms have on the creativity-relevant properties of the whole society or individual agents within it. The analysis can be performed independently from the implemented interaction details making it easy to change the interaction behaviour or properties of the agents and performing the same analysis again to compare the results.

The CSF4CAS provides multiple ways to describe and analyse creative societies with respect to the object-level septuples of individual agents within them and the societal representations $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ and $E_{\mathbf{S}}$. First, it can be used to analyse society $\mathbf{S}$ both on particular time steps t and how the society changes over time. Second, the properties of single agents or some subset of agents may be compared to the society-wide properties on particular time steps, e.g. how well the conceptual space of a single agent, $R_{A_i^t}$, matches that of the whole society's understanding of the conceptual space $R_{\mathbf{S}^t}$. The relationships between these properties may also be observed over time, e.g. an agent can be observed to merge into the society's "norms" if its conceptual space R_{A_i} becomes more similar to the society's $R_{\mathbf{S}}$ on some time interval $[t, t+m]$ where $m > 0$, while the society-wide understanding does not change significantly. Third, the effects of the particular Π implementations may be observed on the societal scale or how they affect its individuals.

Overall, the CSF4CAS provides a novel point of view to analysing creative societies which is grounded on the well-known analysis framework for creative systems and theory of social creativity. It can be used in conjunction with other analysis methods and approaches previously used in computational social sciences and CSC, such as complex systems frameworks where the focus is on interactions between the individual entities, statistical methods which provide quantitative measures of the societies, and multi-agent learning which investigates how learning of multiple individuals can be orchestrated towards a common goal.

4.5 Collaboration

Collaboration is an important aspect of human social creativity [80]. By collaborating with others, humans are able to produce end products that they could not produce by themselves [109]. Collaboration aspects of human behaviour have also been argued to be essential for the initial development of human language [108], which can be understood as a creative behaviour in its own right.

In AI, collaboration has been studied, e.g. in the context of human-robot interaction and collaboration between multiple computational agents. In Computational Creativity collaboration has been studied especially in co-creativity, where a human collaborates with a piece of software [e.g. 53] or an embodied robot [e.g. 46]. Studies that investigate creative collaboration between multiple computational agents are done mainly in the musical domain, where the set of agents which collaborate is fixed during the collaboration, e.g. an ensemble of music bots that improvise a piece of music together [32, 33].

From a more theoretical viewpoint, a true collaboration between computational agents has been argued to be only reachable when the agents have an understanding of each other's *goals* [14]. For example, an agent which sees from a window another agent outside the building has to understand that the other agent's goal includes getting inside the building, after which the door can be opened as a collaboration. If the agent inside opens the door without modelling the other agent's goals, then collaboration did not take place (as understood by the agents), although the other agent is still able to get inside. That is, *true collaboration requires modelling the properties and goals of other agents.* We call these models *peer models.*

In Computational Creativity, McGregor et al. [75] have argued that the goals of the autonomously creative agents should be formed in the interaction between the agent and the environment and other agents. The agents should observe their environment so that they are able to capture interesting (and socially relevant) goals during their execution. Moreover, Kantosalo and Toivonen [52] have presented formal models, based on the CSF (Section 2.3), for two co-creativity process modes between two entities, e.g. a human and software: the *task-divided* co-creativity and *alternating* co-creativity. In task-divided co-creativity, the two entities have different types of tasks they can perform and the task to be performed determines which of the entities proceeds with the co-creative process. In alternating co-creativity, both entities have the same general skills and the co-creative process proceeds by the entities taking the initiative on alternating time steps to develop the co-created end product further.

4.6 Empirical Collaboration Simulations

In this thesis, Paper IV and Paper V study collaboration in iterative simulations in the domain of abstract art. In the simulations, independent artificial agents *learn peer models* of each other and utilise the peer models to *select favourable collaboration partners and socially relevant and interesting goals.* That is, the agents in our setting conform to the (minimal) requirements of intentional collaboration [14] and social goal selection in creative contexts [75].

Paper IV provides the initial simulation, agent and peer model setups, and Paper V is a direct continuation where new peer models are tested and the setup is made more complex by allowing the agents to adjust their aesthetic preferences during the simulation. In general, the agents in the simulations conform to our definition of a creative agent (Definition 3) and use Reinforcement Learning techniques (Section 3.3) to gain interaction-awareness and goal-awareness (Section 2.5).

4.6.1 Overview and Results

In our simulations, a set of creative agents produce images using *genetic programming* [101] and evaluate them using varying *computational aesthetics* [31]. The agents produce images both by themselves and in pairwise collaborations. Our main interest is in how the agents may select beneficial collaboration partners, i.e. partners where both of the agents agree that the collaboration result is creative, and what kind of societal phenomena can be observed when the agents use different collaborator selection methods [Papers IV and V] and when the agents are able to change their aesthetics during the simulation [Paper V]. In other words, our interest in collaboration simulations is in efficient peer modelling (the engineering approach which borrows ideas from the cognitive perspective) and in emergence (social aspects and CSC). To this end, we devise in Paper IV a collaboration procedure for genetic programming where two independent agents may work on the same artefact in a similar type of process as alternating co-creativity [52].

The agents *learn peer models* using simple RL techniques by observing artworks done by themselves and their peers. We investigate whether selfish or altruistic cues of other agents' behaviour are more suitable for learning the peer models for collaborator selection. Selfish cues gauge how much the agent likes its peer's individual artworks, and altruistic cues gauge how well the peers assess the agent's own individual artworks. We conclude that selfish cues seem to give better results, but part of the observed differences may be explained by our collaboration procedure which may reward selfish behaviour more than altruistic behaviour.

The agents use the learned peer models to gain interaction-awareness [Papers IV and V] and goal-awareness [Paper V] which they use in their decision making. Using the peer models the agents may order their peers into collaboration preference order and select the peer which they observe to be the most promising collaboration partner (interaction-awareness, selfish behaviour), and they can adjust their own current aesthetic goal to fit their collaborator partner so that they become "the best possible" partner for it (goal-awareness, altruistic behaviour). We observe, that goal-awareness benefits the collaboration by raising its success ratio and the value (as assessed by the agents involved in the collaboration) of the collaborated artworks.

As the agents (in some learning models) actively inspect what kind of artworks others are producing, i.e. estimate the aesthetics the peers are using, the agents may use the peer models to select a beneficial spot, i.e. an aesthetic goal, within the society. The agents aim to select aesthetic goals which allow them to not only produce novel individual artworks but also, at the same time, find collaboration partners with matching enough interests so that the collaboration may be fruitful (goal-awareness, selfish behaviour with societal impact). We then observe what kind of societal behaviour this strategic adjustment of aesthetic goals brings forth. We conclude that while allowing the agents to adjust their aesthetic goals freely generates more novelty and value, society as a whole behaves more erratically as the agents opportunistically move into similar aesthetic goals in clusters. This is caused by all the agents simultaneously deciding if they want to adjust their aesthetic goal and how they want to do it. An iterative procedure where one agent at a time selects its new aesthetic goal might fix the erratic behaviour, but it would need more controlled societal procedures which are not always desirable or possible (e.g. the designer needs to decide who gets to select their goal first and why).

Altogether, our collaboration simulations provide us with a few key insights. First, what kind of simple peer models can be used to approximate the collaboration potential of other agents. Second, our pairwise collaboration procedure provides the first step towards more informed collaboration procedures between independent agents producing abstract art with genetic programming in social contexts.

Even though our simulation models are not particularly simple (the second requirement of CSC, see Section 4.2), we still relate our work to the general CSC context as our interest is partly on the individual agent properties, i.e. aesthetic goals and peer models, and partly on emergence. As Conte et al. [26] point out and Saunders and Bown [96] sympathise, *it might be that the low hanging fruits of computational social sciences have already been picked up, and the field (and thus CSC) may need to advance towards more complex models and simulations*, which model and capture pervasive aspects of social behaviour, e.g. social creativity. To this end, the empirical simulations done in Paper IV and Paper V, provide the first steps towards substantially more complex social creativity simulations with more ambitious goals.

4.6.2 Analysis of the Simulations

The agents in our empirical cooperation simulations and their societal behaviour may be analysed using the analysis frameworks and tools proposed in this thesis in the following manner.

As already covered, the agents exhibit both interaction-awareness and goal-awareness as proposed in Paper I. Interaction-awareness arises from the agents selecting the most promising collaboration partners. Goal-awareness is based on the agents being able to adjust their goal with respect to their collaboration partner and selecting promising spots within the society to produce novel and valuable artworks, while still being able to find reasonable collaboration partners.

In the Creative Action Selection Framework (CASF), the agents can be seen to follow the general model-free action selection process depicted in Paper II in Algorithm 1. However, the CASF considers only single agents, whereas the agents in our simulations use action selection to select collaboration partners, the actions do not move the agent in the concept or artefact spaces and instead define the peers selected for collaboration. Including interaction actions in the CASF is possible, but it needs elaboration in the framework as the original CSF does not consider interaction with other creative entities or the environment in general.

In the Creative Systems Framework for Creative Agent Societies (CSF4CAS) we consider the strategic movement of our agents in Paper V, in other words socially novel goal selection, to be characterised in the following manner: for each agent A it is likely that for every peer $B \in \mathbf{S} \setminus A$ the set of artefacts which are valued by agent A, E_A, differs considerably from the set of artefacts valued by peer B, E_B. The resulting dynamics caused by the agents adjusting their aesthetic goals can also be characterised on a societal level as the tendency of the set of socially valued artefacts, $E_{\mathbf{S}^t}$, to be in the constant move. However, in our simulations we do not have an explicit process for computing $E_{\mathbf{S}^t}$ as the domain is only implicitly present in the memory of the agents. Thus, we would need to define the process that selects which artefacts are accepted into the domain based on the evaluations of the individual agents.

Chapter 5

Research Task Outcomes

In Chapters 1–4 we have introduced the relevant conceptual elements and perspectives to better understand the original research done in Papers I–V. In this chapter, we present the outcomes of the research tasks as introduced in Section 1.5.

> **Research Task I:** *Design a tool to describe and characterise artificial metacreative systems.*

The first research task is addressed in Paper I and in Section 2.5. We have introduced, based on existing research on self-adaptive and self-aware systems, a set of domain-general self-awareness aspects which are particularly relevant for metacreative systems aiming to produce novel and valuable end products. These self-awareness aspects are artefact-awareness, generator-awareness, goal-awareness, interaction-awareness, time-awareness and meta-self-awareness, each dealing with a particular aspect of a metacreative system's behaviour. With the self-awarenesses, we can provide general characterisations of a system's metacreative capabilities. These characterisations may be compared between systems, even if the systems produce artefacts (or ideas or concepts) in different domains. However, the exact connections between the self-awareness aspects are decided by the designer of the system, which may make the detailed comparisons between systems challenging as reflecting on an aspect may result in control of another aspect.

> **Research Task II:** *Design formal analysis tools for (a) the artefact creation process of individual creative agents and (b) processes in creative societies.*

The second research task is addressed in Paper II and Paper III, and in Section 3.4 and Section 4.4.

Paper II proposes an analysis framework, called Creative Action Selection Framework, for an individual creative agent's action selection by synthesising elements from Markov Decision Processes [5] and their solutions to the Creative Systems Framework [115, 116, 118]. The framework can be used in analysing the action selection capabilities of diverse agents, ranging from simple and minimally intelligent agents to agents which utilise complex learning mechanisms to adjust their action selection process.

Paper III proposes an analysis framework for societies of creative agents, Creative Systems Framework for Creative Agent Societies. The framework introduces elements from the Domain-Individual-Field-Interaction model [30, 35] to the Creative Systems Framework [115, 116, 118]. The framework can be utilised in analysing the overall behaviour of the society, in comparing the properties of individuals to the properties of the whole society, and in investigating how different societal aggregation processes affect the society and its individuals. All of these methods can provide characterisations of creative agent societies and the individuals within them.

Both of the frameworks are domain-general and theoretical.

> **Research Task III:** *Design and implement a society of creative agents to study different collaborator selection procedures and observe phenomena emerging from those procedures.*

The third research task is studied in the context of agents producing abstract art using genetic programming, and it is addressed in Paper IV and Paper V, and in Section 4.6. Using a devised pairwise collaboration procedure we investigate whether selfish or altruistic cues are more reliable in approximating another agent's collaboration potential using simple Reinforcement Learning methods. We conclude that the selfish cues generate more novelty and value, but it may be partly caused by features of our collaboration procedure. Further, we inspect the dynamics of the society where the agents are able to adjust their aesthetic preferences. We observe different emergent society-wide patterns depending on how much the agents are willing to adjust their aesthetics in one go. We envision that modest continuous adjustments may actually increase the stability of the society, as agents may then exploit their accumulated knowledge better.

Chapter 6

Conclusions

We have studied creativity in the context of Computational Creativity, which not only models, replicates and simulates human creativity using machines but also studies how computational entities (systems, agents and societies) may be creative in their own right. We have proposed three different analysis tools or frameworks for creative systems (self-awareness aspects for metacreative systems), agents (Creative Action Selection Framework) and societies (Creative Systems Framework for Creative Agent Societies), and investigated collaboration of creative agents in two case studies in a single creative domain.

Our contributions have added to the understanding of the Producer, Process and Press perspectives to computational creativity. We have (1) proposed relevant conceptual components by which metacreative systems may be characterised (Producer perspective), (2) proposed a framework for analysing action selection within creative agents (Process perspective), and (3) addressed social aspects of creativity and proposed a framework for analysing them (Process and Press perspectives). For the collaboration studies (Process and Press perspectives), we have implemented iterative simulations where independent creative agents produce abstract art by themselves and in pairs and investigated how the agents may learn apt peer models to aid in selecting promising collaboration partners and what kind of societal behaviour emerges.

From a theoretical point of view, the preliminary frameworks proposed in Paper II, the Creative Action Selection Framework (CASF), and in Paper III, the Creative Systems Framework for Creative Agent Societies (CSF4CAS), seem promising to be further developed as general formal frameworks for the analysis of creative agents and societies. In the future their ideas could be combined so that CASF incorporates into itself the interactions between agents, and CSF4CAS adopts the elaborations to the individual agents as proposed in CASF. In addition, CASF would benefit from elaborating the action selection on the meta-level in a similar manner as it is currently done on the object level. This would bring CASF close to the ideas of self-adaptivity, self-awareness and metacreativity as proposed in Paper I, potentially merging them in some form into the framework. In this light, the contributions in this thesis offer one possible way to begin constructing "a unified formal analysis framework" for computational creativity which would take into consideration properties of individual agents, their creative processes and dynamics of their interaction.

To conclude, even though creativity is a multi-faceted phenomenon with domain-specific aspects, it is possible to construct analysis frameworks, and to some extent experiments, where it is handled in a domain-general manner. Nonetheless, there is still a lot to be discovered when it comes to the creativity of artificial entities, as their specific goals, capabilities, architectures, processes and interactions are constantly evolving with every new technological advance we make. This makes it even more crucial to be able to inspect, understand and discuss artificial creative entities apart from their implementation details. Otherwise, the computational creativity field is in danger of scattering into the domain- and technology-specific branches, which do not see the common themes and concepts across them.

References

[1] ADAMATZKY, A., ARMSTRONG, R., JONES, J., AND GUNJI, Y.-P. On creativity of slime mould. *International Journal of General Systems 42*, 5 (2013), 441–457.

[2] ALBERT, R. S., AND RUNCO, M. A. A history of research on creativity. In *Handbook of Creativity*, R. J. Sternberg, Ed. Cambridge University Press, 1998, pp. 16–32.

[3] ALNAJJAR, K., HADAYTULLAH, H., AND TOIVONEN, H. "Talent, skill and support." A method for automatic creation of slogans. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, 2018), Association for Computational Creativity, pp. 88–95.

[4] ALVARADO, J., AND WIGGINS, G. A. Exploring the engagement and reflection model with the creative systems framework. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, 2018), Association of Computational Creativity, pp. 200–207.

[5] BELLMAN, R. A Markovian decision process. *Journal of Mathematics and Mechanics* (1957), 679–684.

[6] BERLYNE, D. E. *Aesthetics and psychobiology.* Appleton-Century-Crofts, 1971.

[7] BODEN, M. A. Creativity and artificial intelligence. *Artificial Intelligence 103*, 1 (1998), 347–356. Artificial Intelligence 40 years later.

[8] BODEN, M. A. *The creative mind: Myths and mechanisms.* Routledge, 2004.

[9] BONABEAU, E. Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences 99*, suppl 3 (2002), 7280–7287.

[10] BOTELLA, M., ZENASNI, F., AND LUBART, T. What are the stages of the creative process? What visual art students are saying. *Frontiers in Psychology 9* (2018), 22–66.

[11] BROCK, A., DONAHUE, J., AND SIMONYAN, K. Large scale GAN training for high fidelity natural image synthesis. In *Proceedings of the Sixth International Conference on Learning Representations* (2018), OpenReview.net.

[12] BUCHANAN, B. G. Creativity at the metalevel: AAAI-2000 presidential address. *AI Magazine 22*, 3 (2001), 13–28.

[13] CARLSON, K., SCHIPHORST, T., AND PASQUIER, P. Scuddle: Generating movement catalysts for computer-aided choreography. In *Proceedings of the Second International Conference on Computational Creativity* (México City, México, 2011), pp. 123–128.

[14] CASTELFRANCHI, C. Modelling social action for AI agents. *Artificial Intelligence 103*, 1 (1998), 157–182.

[15] COLIN, T. R., BELPAEME, T., CANGELOSI, A., AND HEMION, N. Hierarchical reinforcement learning as creative problem solving. *Robotics and Autonomous Systems 86* (2016), 196–206.

[16] COLTON, S. Experiments in meta-theory formation. In *Proceedings of the AISB'01 Symposium on AI and Creativity in Arts and Science* (Heslington, United Kingdom, 2001), pp. 100–109.

[17] COLTON, S. Creativity versus the perception of creativity in computational systems. In *AAAI Spring Symposium: Creative Intelligent Systems* (2008), pp. 14–20.

[18] COLTON, S., BUNDY, A., AND WALSH, T. Agent based cooperative theory formation in pure mathematics. In *Proceedings of the AISB-00 Symposium on Creative & Cultural Aspects and Applications of AI & Cognitive Science* (2000), pp. 11–18.

[19] COLTON, S., CHARNLEY, J., AND PEASE, A. Computational creativity theory: The FACE and IDEA descriptive models. In *Proceedings of the Second International Conference on Computational Creativity* (México City, México, 2011), División de Ciencias de la Comunicación y Diseño Universidad Autónoma Metropolitana – Unidad Cuajimalpa, pp. 90–95.

[20] Colton, S., Cook, M., Hepworth, R., and Pease, A. On acid drops and teardrops: Observer issues in computational creativity. In *Proceedings of the 7th AISB Symposium on Computing and Philosophy* (2014), Society for the Study of Artificial Intelligence and Simulation of Behaviour, pp. 1–8.

[21] Colton, S., Halskov, J., Ventura, D., Gouldstone, I., Cook, M., and Perez-Ferrer, B. The painting fool sees! New projects with the automated painter. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, 2015), pp. 189–196.

[22] Colton, S., Pease, A., Cook, M., and Chen, C. The HR3 system for automatic code generation in creative settings. In *Proceedings of the Tenth International Conference on Computational Creativity* (Charlotte, North Carolina, USA, 2019), Association for Computational Creativity, pp. 108–115.

[23] Colton, S., Pease, A., Corneli, J., Cook, M., and Llano, T. Assessing progress in building autonomously creative systems. In *The Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, 2014), pp. 137–145.

[24] Colton, S., Powley, E. J., and Cook, M. Investigating and automating the creative act of software engineering: A position paper. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, 2018), Association for Computational Creativity, pp. 224–231.

[25] Colton, S., and Wiggins, G. A. Computational creativity: The final frontier? In *Proceedings of the 20th European Conference on Artificial Intelligence* (Amsterdam, The Netherlands, 2012), IOS Press, pp. 21–26.

[26] Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J. P., Sanchez, A., Nowak, A., Flache, A., San Miguel, M., and Helbing, D. Manifesto of computational social science. *The European Physical Journal Special Topics 214*, 1 (Nov. 2012), 325–346.

[27] Cook, M., and Colton, S. Generating code for expressing simple preferences: Moving on from hardcoding and randomness. In

Proceedings of the Sixth International Conference on Computational Creativity (Park City, Utah, 2015), Brigham Young University, pp. 8–16.

[28] COOK, M., COLTON, S., AND GOW, J. The ANGELINA videogame design system – part II. *IEEE Transactions on Computational Intelligence and AI in Games 9*, 3 (2017), 254–266.

[29] CRNKOVIC-FRIIS, L., AND CRNKOVIC-FRIIS, L. Generative choreography using deep learning. In *Proceedings of the Seventh International Conference on Computational Creativity* (Paris, France, 2016), Sony CSL, pp. 272–277.

[30] CSIKSZENTMIHALYI, M. Society, culture, and person: A systems view of creativity. In *The Nature of Creativity: Contemporary Psychological Perspectives*, R. J. Sternberg, Ed. Cambridge University Press, 1988, pp. 325–339.

[31] DEN HEIJER, E., AND EIBEN, A. Investigating aesthetic measures for unsupervised evolutionary art. *Swarm and Evolutionary Computation 16* (2014), 52–68.

[32] EIGENFELDT, A., BOWN, O., BROWN, A. R., AND GIFFORD, T. Distributed musical decision-making in an ensemble of musebots: Dramatic changes and endings. In *Proceedings of the Eight International Conference on Computational Creativity* (Atlanta, Georgia, USA, 2017), Georgia Institute of Technology, pp. 88–95.

[33] EIGENFELDT, A., BOWN, O., AND CASEY, B. Collaborative composition with creative systems: Reflections on the first musebot ensemble. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, 2015), Brigham Young University, pp. 134–141.

[34] EPSTEIN, J. M., AND AXTELL, R. *Growing artificial societies: Social science from the bottom up.* Brookings Institution Press, 1996.

[35] FELDMAN, D., CSIKSZENTMIHALYI, M., AND GARDNER, H. *Changing the world: A framework for the study of creativity.* Praeger, Westport, CT, 1994.

[36] GABORA, L. Honing theory: A complex systems framework for creativity. *Nonlinear dynamics, psychology, and life sciences 21*, 1 (2017), 35–88.

[37] Gabora, L., and Tseng, S. The social impact of self-regulated creativity on the evolution of simple versus complex creative ideas. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, 2014), Jožef Stefan Institute, pp. 8–15.

[38] Garlan, D., Cheng, S.-W., Huang, A.-C., Schmerl, B., and Steenkiste, P. Rainbow: Architecture-based self-adaptation with reusable infrastructure. *Computer 37*, 10 (Oct. 2004), 46–54.

[39] Glines, P., Griffith, I., and Bodily, P. Software design patterns of computational creativity: A systematic mapping study. In *Proceedings of the Twelfth International Conference on Computational Creativity* (México, 2021), Association for Computational Creativity, pp. 218–221.

[40] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in Neural Information Processing Systems 27* (2014), Curran Associates, Inc., pp. 2672–2680.

[41] Grace, K., and Maher, M. L. Specific curiosity as a cause and consequence of transformational creativity. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, 2015), Brigham Young University, pp. 260–267.

[42] Guckelsberger, C., Salge, C., and Togelius, J. New and surprising ways to be mean. In *IEEE Conference on Computational Intelligence and Games (CIG)* (Maastricht, Netherlands, 2018), IEEE, pp. 1–8.

[43] Hadjeres, G., Pachet, F., and Nielsen, F. Deepbach: A steerable model for bach chorales generation. In *Proceedings of the 34th International Conference on Machine Learning - Volume 70* (2017), JMLR.org, pp. 1362–1371.

[44] Hantula, O., and Linkola, S. Towards goal-aware collaboration in artistic agent societies. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, 2018), Association of Computational Creativity, pp. 136–143.

[45] Hekkert, P., and Van Wieringen, P. C. W. Beauty in the eye of expert and nonexpert beholders: A study in the appraisal of art. *The American Journal of Psychology 109*, 3 (1996), 389–407.

[46] Hoffman, G., and Weinberg, G. Gesture-based human-robot jazz improvisation. In *Proceedings of the IEEE International Conference on Robotics and Automation* (2010), IEEE, pp. 582–587.

[47] Hutchings, P., and McCormack, J. Using autonomous agents to improvise music compositions in real-time. In *Proceedings for the Sixth International Conference on Computational Intelligence in Music, Sound, Art and Design* (Amsterdam, The Netherlands, 2017), Springer International Publishing, pp. 114–127.

[48] Jaworski, W. Mind-body theories and mind-body problems. In *Philosophy of Mind: A Comprehensive Introduction*. Wiley-Blackwell, 2011, pp. 1–22.

[49] Jennings, K. E. Developing creativity: Artificial barriers in artificial intelligence. *Minds and Machines 20*, 4 (2010), 489–501.

[50] Jordanous, A. A standardised procedure for evaluating creative systems: Computational creativity evaluation based on what it is to be creative. *Cognitive Computation 4*, 3 (Sept. 2012), 246–279.

[51] Jordanous, A. Four PPPPerspectives on computational creativity in theory and in practice. *Connection Science 28*, 2 (2016), 194–216.

[52] Kantosalo, A., and Toivonen, H. Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In *Proceedings of the Seventh International Conference on Computational Creativity* (Paris, France, 2016), Sony CSL, pp. 77–84.

[53] Kantosalo, A. A., Toivanen, J. M., and Toivonen, H. T. T. Interaction evaluation for human-computer co-creativity: A case study. In *Proceedings of the Sixth International Conference on Computational Creativity* (Provo, Utah, USA, 2015), Brigham Young University, pp. 276–283.

[54] Kaufman, A. B., and Kaufman, J. C. *Animal creativity and innovation*. Academic Press, 2015.

[55] Kephart, J. O., and Chess, D. M. The vision of autonomic computing. *Computer 36*, 1 (Jan. 2003), 41–50.

[56] Kounev, S., Kephart, J., Milenkoski, A., and Zhu, X. *Self-Aware Computing Systems*. Springer International Publishing, 2017.

[57] Kozbelt, A., Beghetto, R. A., and Runco, M. A. Theories of creativity. In *The Cambridge Handbook of Creativity*, J. C. Kaufman and R. J. Sternberg, Eds. Cambridge University Press, 2010, pp. 20–47.

[58] Laddaga, R. DARPA broad agency announcement on self-adaptive software (BAA-98-12), 1997.

[59] Lamb, C., Brown, D. G., and Clarke, C. Human competence in creativity evaluation. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, USA, 2015), Brigham Young University, pp. 102–109.

[60] Lamb, C., Brown, D. G., and Clarke, C. L. A. Evaluating computational creativity: An interdisciplinary tutorial. *ACM Computing Surveys 51*, 2 (Feb. 2018), 1–34.

[61] Lewis, P. R., Chandra, A., Faniyi, F., Glette, K., Chen, T., Bahsoon, R., Torresen, J., and Yao, X. Architectural aspects of self-aware and self-expressive computing systems: From psychology to engineering. *Computer 48*, 8 (2015), 62–70.

[62] Liapis, A., Yannakakis, G. N., and Togelius, J. Computational game creativity. In *Proceedings of the Fifth International Conference on Computational Creativity* (Ljubljana, Slovenia, 2014), pp. 46–53.

[63] Linkola, S., Guckelsberger, C., and Kantosalo, A. Action selection in the creative systems framework. In *Proceedings of the Eleventh International Conference on Computational Creativity* (Coimbra, Portugal, 2020), Association for Computational Creativity, pp. 303–310.

[64] Linkola, S., and Hantula, O. On collaborator selection in creative agent societies: An evolutionary art case study. In *Proceedings for the Seventh International Conference on Computational Intelligence in Music, Sound, Art and Design* (Parma, Italy, 2018), Springer International Publishing, pp. 206–222.

[65] Linkola, S., and Kantosalo, A. Extending the creative systems framework for the analysis of creative agent societies. In *Proceedings of the Tenth International Conference on Computational Creativity* (Charlotte, NC, USA, 2019), Association for Computational Creativity, pp. 204–211.

[66] LINKOLA, S., KANTOSALO, A., MÄNNISTÖ, T., AND TOIVONEN, H. Aspects of self-awareness: An anatomy of metacreative systems. In *Proceedings of the Eight International Conference on Computational Creativity* (Atlanta, Georgia, USA, 2017), Georgia Institute of Technology, pp. 189–196.

[67] LINKOLA, S., MÄKITALO, N., AND MÄNNISTÖ, T. On the inherent creativity of self-adaptive systems. In *Proceedings of the Eleventh International Conference on Computational Creativity* (Coimbra, Portugal, 2020), Association for Computational Creativity, pp. 362–365.

[68] LINKOLA, S., TAKALA, T., AND TOIVONEN, H. Novelty-seeking multi-agent systems. In *Proceedings of the Seventh International Conference on Computational Creativity* (Paris, France, 2016), Sony CSL, pp. 1–8.

[69] LIU, Y.-T. Creativity or novelty?: Cognitive-computational versus social-cultural. *Design Studies 21*, 3 (2000), 261–276.

[70] LOUGHRAN, R., AND O'NEILL, M. Application domains considered in computational creativity. In *Proceedings of the Eighth International Conference on Computational Creativity* (Atlanta, GA, 2017), Association for Computational Creativity, pp. 197–204.

[71] LOUGHRAN, R., AND O'NEILL, M. Is computational creativity domain-general? In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, 2018), Association for Computational Creativity, pp. 112–119.

[72] LUBART, T. I. Models of the creative process: Past, present and future. *Creativity Research Journal 13*, 3-4 (2001), 295–308.

[73] MACHADO, P., AND CORREIA, J. A. Semantic aware methods for evolutionary art. In *Proceedings of the 2014 Annual Conference on Genetic and Evolutionary Computation* (New York, NY, USA, 2014), ACM, pp. 301–308.

[74] MACKINNON, D. W. Creativity: A multi-faceted phenomenon. In *Creativity: A discussion at the nobel conference* (1970), North-Holland Publishing Company, pp. 17–32.

[75] MCGREGOR, S., MCGINITY, M. M., AND GRIFFITHS, S. How many robots does it take? Creativity, robots and multi-agent systems. In *The AISB15's Second International Symposium on Computational Creativity* (2015), pp. 23–29.

[76] Minsky, M. L. Why people think computers can't. *AI magazine 3*, 4 (1982), 3–15.

[77] Mumford, M., and Ventura, D. The man behind the curtain: Overcoming skepticism about creative computing. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, 2015), Brigham Young University, pp. 1–7.

[78] Newell, A., Shaw, J. C., and Simon, H. A. The processes of creative thinking. In *Contemporary Approaches to Creative Thinking* (1962), Atherton Press.

[79] Pasquier, P., Eigenfeldt, A., Bown, O., and Dubnov, S. An introduction to musical metacreation. *Computers in Entertainment 14*, 2 (Jan. 2017), 2:1–2:14.

[80] Paulus, P., and Nijstad, B. *Group Creativity: Innovation through Collaboration*. Oxford University Press, 2003.

[81] Pease, A., and Colton, S. Computational creativity theory: Inspirations behind the FACE and the IDEA models. In *Proceedings of the Second International Conference on Computational Creativity* (México City, México, Apr. 2011), División de Ciencias de la Comunicación y Diseño Universidad Autónoma Metropolitana – Unidad Cuajimalpa, pp. 72–77.

[82] Pérez y Pérez, R. The computational creativity continuum. In *Proceedings of the Ninth International Conference on Computational Creativity* (Salamanca, Spain, 2018), Association for Computational Creativity, pp. 177–184.

[83] Pérez y Pérez, R., Negrete, S., Nalosa, E. P., Ávila, R., Castellanos, V., and Lemaitre, C. MEXICA-Impro: A computational model for narrative improvisation. In *Proceedings of the International Conference on Computational Creativity* (Lisbon, Portugal, 2010), Department of Informatics Engineering, University of Coimbra, pp. 90–99.

[84] Pérez y Pérez, R., and Sharples, M. MEXICA: A computer model of a cognitive account of creative writing. *Journal of Experimental & Theoretical Artificial Intelligence 13*, 2 (2001), 119–139.

[85] Pettee, M., Shimmin, C., Douglas, D., and Vidrin, I. Beyond imitation: Generative and variational choreography via machine

learning. In *Proceedings of the Tenth International Conference on Computational Creativity* (Charlotte, North Carolina, USA, 2019), Association for Computational Creativity, pp. 196–2003.

[86] PLUCKER, J., AND BEGHETTO, R. A. Why creativity is domain general, why it looks domain specific, and why the distinction does not matter. In *Creativity: From potential to realization.* American Psychological Association, 2004, pp. 153—-167.

[87] PUTERMAN, M. L. *Markov Decision Processes: Discrete Stochastic Dynamic Programming.* Wiley, 2014.

[88] RHODES, M. An analysis of creativity. *The Phi Delta Kappan 42*, 7 (1961), 305–310.

[89] RITCHIE, G. Some empirical criteria for attributing creativity to a computer program. *Minds and Machines 17* (2007), 76–99.

[90] ROZANSKI, N., AND WOODS, E. *Software systems architecture: Working with stakeholders using viewpoints and perspectives.* Addison-Wesley, 2012.

[91] RUMMERY, G. A., AND NIRANJAN, M. On-line Q-learning using connectionist systems. Tech. rep., University of Cambridge, Department of Engineering, Cambridge, England, 1994.

[92] RUNCO, M. A., AND JAEGER, G. J. The standard definition of creativity. *Creativity Research Journal 24*, 1 (2012), 92–96.

[93] RUSSELL, S., AND NORVIG, P. *Artificial Intelligence: A Modern Approach*, 3rd ed. Prentice Hall Press, USA, 2009.

[94] SADLER-SMITH, E. Wallas' four-stage model of the creative process: More than meets the eye? *Creativity Research Journal 27*, 4 (2015), 342–352.

[95] SALEHIE, M., AND TAHVILDARI, L. Self-adaptive software: Landscape and research challenges. *ACM Transactions on Autonomous and Adaptive Systems 4*, 2 (May 2009), 14:1–14:42.

[96] SAUNDERS, R., AND BOWN, O. Computational social creativity. *Artificial Life 21*, 3 (2015), 366–378.

[97] SAUNDERS, R., AND GERO, J. S. Artificial creativity: A synthetic approach to the study of creative behaviour. In *Computational and*

Cognitive Models of Creative Design V (Sydney, Australia, 2001), University of Sydney, University of Sydney, pp. 113–139.

[98] Saunders, R., and Gero, J. S. A curious design agent: A computational model of novelty-seeking behaviour in design. In *Proceedings of the Sixth Conference on Computer Aided Architectural Design Research in Asia* (Sydney, Australia, 2001), vol. 1, pp. 345–350.

[99] Schmidhuber, J. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development 2*, 3 (Sept. 2010), 230–247.

[100] Shoham, Y., and Leyton-Brown, K. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations.* Cambridge University Press, 2008.

[101] Sims, K. Artificial evolution for computer graphics. In *Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 1991), ACM, pp. 319–328.

[102] Sosa, R., and Gero, J. S. Social models of creativity. In *Proceedings of the International Conference of Computational and Cognitive Models of Creative Design VI* (Heron Island, Australia, 2005), Key Centre of Design Computing and Cognition, University of Sydney, Australia, pp. 19–44.

[103] Stein, M. I. Creativity and culture. *The Journal of Psychology 36*, 2 (1953), 311–322.

[104] Sternberg, R., Ed. *Handbook of Creativity.* Cambridge University Press, 1998.

[105] Sternberg, R. J., and Lubart, T. I. The concept of creativity: Prospects and paradigms. In *Handbook of Creativity*, R. J. Sternberg, Ed. Cambridge University Press, 1998, pp. 3–15.

[106] Sutton, R. S., and Barto, A. G. *Reinforcement learning: An introduction.* MIT press, 2018.

[107] Toivonen, H., and Gross, O. Data mining and machine learning in computational creativity. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 5*, 6 (2015), 265–275.

[108] Tomasello, M. *Origins of human communication.* MIT press, 2010.

[109] UZZI, B., AND SPIRO, J. Collaboration and creativity: The small world problem. *American Journal of Sociology 111*, 2 (2005), 447–504.

[110] VENTURA, D. Mere generation: Essential barometer or dated concept? In *Proceedings of the Seventh International Conference on Computational Creativity* (Paris, France, 2016), Sony CSL, pp. 17–24.

[111] VENTURA, D. How to build a CC system. In *Eighth International Conference on Computational Creativity* (Atlanta, GA, USA, 2017), Association for Computational Creativity, pp. 253–260.

[112] VIGORITO, C. M., AND BARTO, A. G. Hierarchical representations of behavior for efficient creative search. In *AAAI Spring Symposium: Creative Intelligent Systems* (2008), The AAAI Press, pp. 135–141.

[113] WATKINS, C. J., AND DAYAN, P. Q-learning. *Machine learning 8*, 3-4 (1992), 279–292.

[114] WEDYAN, F., AND ABUFAKHER, S. Impact of design patterns on software quality: A systematic literature review. *IET Software 14*, 1 (2020), 1–17.

[115] WIGGINS, G. A. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems 19*, 7 (Nov. 2006), 449–458.

[116] WIGGINS, G. A. Searching for computational creativity. *New Generation Computing 24*, 3 (Sept. 2006), 209–222.

[117] WIGGINS, G. A. The mind's chorus: Creativity before consciousness. *Cognitive Computation 4*, 3 (2012), 306–319.

[118] WIGGINS, G. A. A framework for description, analysis and comparison of creative systems. In *Computational Creativity*. Springer, 2019, pp. 21–47.

[119] WIGGINS, G. A. Creativity, information, and consciousness: The information dynamics of thinking. *Physics of life reviews 34* (2020), 1–39.

[120] XIAO, P., ALNAJJAR, K., GRANROTH-WILDING, M., AGRES, K., AND TOIVONEN, H. Meta4meaning: Automatic metaphor interpretation using corpus-derived word associations. In *Proceedings of*

the Seventh International Conference on Computational Creativity (Paris, France, 2016), Sony CSL, pp. 230–237.

[121] Xiao, P., and Linkola, S. Vismantic: Meaning-making with images. In *Proceedings of the Sixth International Conference on Computational Creativity* (Park City, Utah, June 2015), Brigham Young University, pp. 158–165.