Extending the Creative Systems Framework for the Analysis of Creative Agent Societies

Simo Linkola and Anna Kantosalo
Department of Computer Science
University of Helsinki
{simo.linkola,anna.kantosalo}@helsinki.fi

Abstract

Design and analysis of creative agent societies is often done in the context of computational sociology, which has evolved into its own field called computational social creativity (Saunders and Bown 2015). In this paper, we propose a formal framework for describing and analysing creative agent societies based on the Creative Systems Framework (CSF) (Wiggins 2006a; 2006b). We extend the CSF for single agents to include creative agent societies. The extended CSF allows us to describe society wide phenomena relevant for creativity, identify how individual agents relate to the whole society and characterise societal consequences caused by adopting certain policies. All these formal tools may be used when analysing designated creative agent societies. We demonstrate this by devising a straightforward practical procedure which may be used to gain insight into the influence individual agents have on the society over time.

Introduction

The Creative Systems Framework (CSF) (Wiggins 2006a; 2006b), defines exploratory (and transformational) creativity (Boden 1992) with mathematical rigour. It provides tools for describing and analysing diverse interesting phenomena, which may occur in creative systems. In this paper, we propose a minimal extension to the CSF, allowing us to utilise the CSF's formalisms to conduct extensive analysis on creative agent societies where each agent exhibits exploratory creativity. We call this extension the CSF for creative agent societies.

Computational social creativity (CSC) (Saunders and Bown 2015) is often adopted as a general context when analysing creative agent societies, i.e. multiple creative systems interacting with each other. In CSC, models should (1) demonstrate a mechanism, (2) be simple and reproducible, and (3) preferably generate new hypotheses (Saunders and Bown 2015). By a mechanism we mean that some properties of the individual systems or their interaction provoke observable emergent behaviour in the whole agent society, bringing into focus how various policies regarding artefact exchange and evaluation affect the society. Thus, demonstrating a mechanism implies that the preferred method used to analyse the society is of qualitative nature.

A prominent theoretical background used in CSC is Csikszentmihalyi's systems view of creativity (Csikszentmihalyi 1988), which has later been developed into the Domain-Individual-Field Interaction (DIFI) model. Its main argument is that creativity is not isolated in any individual. Instead, creativity can be understood in its entirety only by studying dynamic interactions between *individual* producers in the society, experts of a given *field*, and a *domain* of accumulated cultural artefacts (filtered by the field). This view can be seen as an encompassing conceptualisation of the creative process in a society, which itself allows elaborating on any of the following perspectives: what kind of artefacts are accepted into the domain, how a field is formed and how it operates, what are the abilities of the individuals, and how they create.

The Creative Systems Framework is a potent tool for describing and analysing individual creative systems exhibiting exploratory creativity. The CSF has been used as a foundation in several other studies (Grace and Maher 2015; Kantosalo and Toivonen 2016; Alvarado and Wiggins 2018). As a high level framework, the CSF does not take a stance on exactly *how* the exploratory creative process is executed. Rather, its formalisms allow describing and analysing the overall *capabilities* of a creative system's exploratory process, making it possible to identify different phenomena relating to these capabilities.

This paper expands the CSF's mathematical analysis tools into the context of CSC and the DIFI model. The extension to the CSF is designed for two purposes. First, it defines concepts which make it possible to discuss the capabilities of exploratory creative agent societies in detail. Second, it is a conceptual tool for analysing how changes in the properties of the individuals or the field affect the encompassing dynamics of the exploratory creative agent society. As such, it provides a novel point of view to CSC for inspecting if a particular mechanism exists or how the mechanism affects the whole society. The extension can be also utilised to gain insight into what kind of impact individuals or groups of individuals have on the field (e.g. indirect analysis on individuals' role in the society), or how the society's creative capabilities change over time (e.g. how the implemented interactions between the field and the individuals affect the behaviour of the whole society).

The rest of the paper is organised as follows. We begin

$\mathcal{U}$	the universe of all possible concepts
$\mathcal{L}$	a language in which to express concepts and rules
$[\![.]\!]$	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, which maps three subsets of $\mathcal{L}$ to a function that generates a new sequence of elements of $\mathcal{U}$ from an existing one
$\mathcal{R}$	rules defining valid concepts
$\mathcal{T}$	rules defining traversal in the concept space
$\mathcal{E}$	rules defining evaluation of concepts

Table 1: Descriptions of different elements of the CSF.

by providing a brief introduction to the CSF, after which we move on to our contributions. First, we introduce our extension to the CSF, the CSF for creative agent societies, where the nucleus is an abstract societal aggregation function which outputs society wide representations of the CSF's formalisms. Second, we show how the societal aggregation function and the society wide representations can be used to describe and identify various phenomena, both at a single time step and evolving in time, which may occur in creative agent societies. Lastly, we demonstrate that it is possible to derive practical procedures to test hypotheses considering the impact individual agents, or agent subsocieties, have on the whole society and its evolution. The paper ends with discussion and conclusions.

The CSF in a Nutshell

The Creative System's Framework (CSF) (Wiggins 2006a; 2006b) is a formalisation of Boden's (1992) exploratory and transformational creativity. To keep the scope of the paper reasonable, we will concentrate on exploratory creativity. However, we acknowledge that transformational capabilities of individuals are a major cause of emergent phenomena at the societal level, e.g. the field changes based on the changes in the individuals forming it or interacting with it.

According to Wiggins (2006a; 2006b), the exploratory part of the CSF is a septuple

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

The individual elements of the septuple are defined in Table 1. For full definitions, we refer to Wiggins (2006a; 2006b). The universe $\mathcal{U}$ defines all possible (partial) concepts, i.e. artefacts, and a language $\mathcal{L}$ provides means to express those (partial) concepts. In this paper we concentrate on $\mathcal{R}$, $\mathcal{T}$ and $\mathcal{E}$, which are elaborated below.

$\mathcal{R}$ is a set of rules (expressed in $\mathcal{L}$) for selecting acceptable artefacts. Applying a selector function generated from $\mathcal{R}$ by the interpretation function $[\![.]\!]$ (and a proper threshold $p \in [0, 1]$), gives us a way to define which artefacts are considered acceptable or valid by the agent (e.g., art, poetry, etc.). This formulation yields Boden's (1992) conceptual space. Drawing from Kantosalo and Toivonen (2016), we denote the valid subset of concepts in the universe as

$$R \equiv \{c \mid c \in \mathcal{U} \wedge [\![\mathcal{R}]\!](c) \geq p\}.$$

$\mathcal{T}$ is a set of rules (expressed in $\mathcal{L}$), which realise, when interpreted, the traversal behaviour of the agent in the conceptual space. That is, it gives the means for the agent to move from known (even empty) concepts to unknown ones. The traversal may be informed by $\mathcal{R}$ or $\mathcal{E}$, hence the interpretation function $\langle\!\langle ., ., . \rangle\!\rangle$ is used to interpret the rules as behaviour. That is, given a (possibly empty) concept c_{in}, $\langle\!\langle ., ., . \rangle\!\rangle$ outputs a new concept: $c_{\text{out}} = \langle\!\langle \mathcal{R}, \mathcal{T}, \mathcal{E} \rangle\!\rangle(c_{\text{in}})$. In general c_{in} and c_{out} can also be sets of concepts. We follow Kantosalo and Toivonen (2016) and denote the concepts reachable from a given concept c_{in} using at most n recursive steps as $T^n(c_{\text{in}})$:

$$T^n(c_{\text{in}}) \equiv \bigcup_{j=0}^{n} \langle\!\langle \mathcal{R}, \mathcal{T}, \mathcal{E} \rangle\!\rangle^j(c_{\text{in}}),$$

leaving out the superscript n when it is not required. Traversal behaviour defined in this manner is much alike the standard AI search framework.

$\mathcal{E}$ is a set of rules (expressed in $\mathcal{L}$), which define the evaluation of the creative outputs, appropriately contextualised. This contextualisation might be subjective to the agent or some objective comparison. Similarly to conceptual space R, a set of *valued concepts* E may be defined by applying a selector function generated from $\mathcal{E}$ by $[\![.]\!]$ (and a proper threshold $p \in [0, 1]$):

$$E \equiv \{v \mid v \in \mathcal{U} \wedge [\![\mathcal{E}]\!](v) \geq p\}.$$

To summarise, $\mathcal{R}$ defines what kind of artefacts are accepted, $\mathcal{E}$ defines which artefacts are valued, and $\mathcal{T}$ provides the means to explore the artefact space. A prominent aspect which separates the CSF from standard AI search is the use of both $\mathcal{R}$ and $\mathcal{E}$. For example, an independent creator may value slightly different artefacts than which it perceives as acceptable, creating "tension" between $\mathcal{R}$ and $\mathcal{E}$. With proper abilities of the agent, this tension may cause transformation in one of them, changing the exploration process in a fundamental manner. Especially transformation of $\mathcal{R}$, i.e. what types of artefacts are considered as acceptable solutions, is not present in standard AI search. This tension allows Wiggins to identify two types of interesting phenomena, namely *uninspiration* and *aberration* (Wiggins 2006a) driving transformations in the system.

Next, we will move on to describe our extension to the CSF, the CSF for creative agent societies, and how it may be used to describe interesting phenomena, similar to aberration and uninspiration, on a societal level.

The CSF for Creative Agent Societies

In this section we define our extension to the CSF. We begin by covering basic notation and assumptions. Then, we show how the CSF is interpreted for single agents with the addition of the input parameters presenting information exchange in the society, after which we move on to describe the societal elements of our extension: the societal aggregation function Π and the societal R, T and E.

The proposed extension is not "complete" in the manner of the CSF, as we do not provide a full formalism for creative agent societies. Instead, we present the minimal elements required to (1) discuss exploratory creativity on a societal level and (2) describe and analyse creative agent societies.

Our extension has two principal assumptions considering the agents in the creative society. First, each agent is considered as *an independent creator*. While the other agents may have an influence over the agent, ultimately the agent itself is in charge of its own creative process. Second, each agent in the society is assumed to exhibit exploratory creativity. This type of creativity is commonly used in creative agent societies, and it can be implemented with diverse generative methods. For example, genetic algorithms, many deep learning models and several other methods under the generate-and-test paradigm (Toivonen and Gross 2015) can be harnessed to obtain exploratory creativity. However, despite these design characteristics, many of the analysis tools presented in this paper can be applied with careful consideration to societies composed of other kinds of agents.

We define a society $\mathbf{S}$ as a set of agents, $\mathbf{S} = \{A_1, A_2, \dots, A_n\}$, our standpoint being that inter-agent relationships, norms and other social phenomena are fundamentally (encapsulated in) the properties of the agents. Even though there might be some external representations of these concepts, the agents may (mis)interpret them. Thus, the state of the society and how it operates is functionally described by taking into account each agent's own view of the social structures.

We denote a time step the society $\mathbf{S}$ is going through as t, $t \geq 0$, where the society (or a part of it) is initialised at $t = 0$. The semantics of the time (e.g. is the simulated time continuous or discrete) may vary between actual agent society implementations, however, the formulations in this paper are independent of them. Thus, the society $\mathbf{S}$ on a time step t, $\mathbf{S}^t$, is a set of agents

$$\mathbf{S}^t = \{A_1^t, A_2^t, \dots, A_n^t\}.$$

We omit the superscript notation specifying the time step t when it is not necessary.

We assume that each agent $A_i \in \mathbf{S}$ is an independent creator with its own private CSF, meaning that the agent itself controls which artefacts it creates. Other agents may have an impact on the agent's creative process, but only the agent itself executes it. We denote the CSF of agent A_i on the time step t as

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

assuming that the first four elements are the same for each agent, i.e. the agents use the same language to describe the concepts in the universe, and have the same interpreter functions. The agents only differ in what they consider as acceptable artefacts, the rules of their traversal behaviour and how they evaluate artefacts.

The agents need to be able to communicate with each other in a society. In creative agent societies, agents typically communicate by exchanging artefacts and the artefact producer's identity may affect the perception of the artefact. Thus, we need to alter the interpreter functions $[\![.]\!]$ and $\langle\langle ., ., . \rangle\rangle$ (as agents may also start their search from (partial) artefacts given to them by their peers) to handle this requirement. We extend both $[\![.]\!]$ and $\langle\langle ., ., . \rangle\rangle$ to accept a second input argument, the identity (or a set of identities) of the agent which produced the artefact, I. As a consequence, agent's A_i conceptual space R_{A_i} is defined as

$$R_{A_i} \equiv \{c \mid c \in \mathcal{U} \wedge [\![\mathcal{R}_{A_i}]\!](c, I) \geq p, \text{ where } I \subseteq \mathbf{S}\},$$

E_{A_i} is handled in the same way, and $T^n_{A_i}$ is defined as

$$T^n_{A_i}(c_{\text{in}}, I) \equiv \bigcup_{j=0}^{n} \langle\langle \mathcal{R}_{A_i}, \mathcal{T}_{A_i}, \mathcal{E}_{A_i} \rangle\rangle^j (c_{\text{in}}, I).$$

We assume that all other communication between the agents (other metadata related to the artefacts, such as framing, communication about beliefs and intentions, etc.) is handled through some other properties or processes of the agents, which are not explicitly presented in the extension. This communication may, however, drive the change in the agent's own creative process, thus transforming some elements of its CSF, e.g. E_A or T_A.

We denote by $R_{A_i^t}$ the set of artefacts considered as valid by A_i on time step t, and by $E_{A_i^t}$ the set of artefacts considered valuable. For creativity, it is especially interesting, when these sets are subject to alterations over time. These alterations can be due to the change in the set of rules an agent uses or the threshold p applied to the outputs of the function interpreted from the rules, e.g. $[\![\mathcal{R}_{A_i}]\!]$.

Π: The Societal Aggregation Function

In the centre of the proposed extension is a societal aggregation function Π, which interprets the properties of the individual agents (their CSFs, relationships, etc.) into society wide concepts, taking into account the different social norms, processes and policies present in the society. By abstracting the interaction between the agents, their relationships, etc., into Π, we lose a lot of meaningful information about a particular society, but gain the elegance of the CSF in describing interesting situations which may arise in it.

Formally, the societal aggregation function Π takes as an argument a set of agents at time step t, $\mathbf{S}^t$, and outputs societal R, T and E for that time step. That is,

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

where the societal structures $R_{\mathbf{S}^t}$, $T_{\mathbf{S}^t}$ and $E_{\mathbf{S}^t}$ are an aggregation of all the individual agents' properties at time step t affected by $\mathbf{S}$'s policies and other societal structures. In the special case when $\mathbf{S} = \{A\}$, i.e. the society contains only a single agent, Π returns the agent's own R_A, T_A and E_A. The societal aggregation function Π acts as an interpreter function for the individual agents and their interactions. Its outputs, $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ and $E_{\mathbf{S}}$, define the (implicit or explicit) social interpretations of which artefacts society perceives as valid, what artefacts can be reached by the society, and which artefacts are valued by the society, respectively.[1]

[1]The fact that Π returns the interpreted $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ and $E_{\mathbf{S}}$ and not the societal rules $\mathcal{R}_{\mathbf{S}}$, $\mathcal{T}_{\mathbf{S}}$ and $\mathcal{E}_{\mathbf{S}}$ still requiring interpretation is purely a design decision. One could also formulate Π to return the latter and define the societal interpretation function for each set of rules. Most of the analysis tools described later in the paper could be very well applied to the sets of rules instead of the sets of artefacts, providing yet another angle to the creative agent society's operation. However, hierarchical societies could benefit from first composing the societal rules and then their interpretations.

Next, we characterise $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ and $E_{\mathbf{S}}$ and how they relate to the DIFI model's concepts. In the following, we will assume that the society has only one field, and all agents may be part of it. That is, the agents in the society do not form isolated subsocieties and they produce concepts within the same domain, e.g. art.

$R_{\mathbf{S}}$: The Societal R for S

$R_{\mathbf{S}}$ is the society $\mathbf{S}$'s collective understanding of which artefacts are accepted as valid, e.g. works of art. Transforming the individual agent properties into $R_{\mathbf{S}}$ has to take into account how the society's policies, communication structures, etc., aggregate individual agents' views of what is accepted as a societal norm of art. However, as our goal is to be able to describe the capabilities of the creative agent society, Π is designed to abstract away exactly the decisions related to how $R_{\mathbf{S}}$ should be formed.

The aggregation of individual properties into societal properties may be implicit or explicit. In the implicit case there is no clearly defined society wide decision procedure for acceptance, whereas in the explicit case there is a procedure (to which agents more or less conform to) which outputs acceptance for each artefact or a fuzzy assignment for it. In the explicit case, $R_{\mathbf{S}}$ is the output of the procedure if it would be run for all artefacts. However, as this is not computationally feasible, typically only the artefacts perceived by the society on a given time step are put into test for inclusion. In the implicit case, the analyst should define a proper way to compute $R_{\mathbf{S}}$ from the properties of the individual agents. From an agent's perspective, the characteristic difference between the two is that the output of the explicit aggregation (given the input) is known to the agent, while in the implicit case the agent needs to model the process by some means.

In practice, the explicit decision procedure might be, e.g., some aggregate measure taking into account each agent's individual R_A and voting for inclusion in $R_{\mathbf{S}}$ for each candidate artefact. Another way would be to compute each agents $[\![\mathcal{R}_A]\!]$ for each artefact, producer identity pair and applying some aggregate function (such as $\min$, $\max$ or mean) to it, which is then thresholded to filter artefacts accepted to $R_{\mathbf{S}}$.

From the DIFI model's point of view, $R_{\mathbf{S}}$ can be seen to be formed by the field. On the other hand, a set of agents with reasonably similar R, which differs from $R_{\mathbf{S}}$, may form their own field, specialising to a particular type of artefacts. The experts in the field define which artefacts are seen as art and, thus, are candidates for domain inclusion. However, this process is dynamic. First, the domain affects the perception of new artefacts, each new artefact included in the domain has an impact on the perception of subsequent artefacts. Second, the identity of individuals producing artefacts affect their social evaluation. Third, the individuals which form the field may change, thus affecting the process.

$T_{\mathbf{S}}$: The Societal T for S

$T_{\mathbf{S}}$ encapsulates the artefacts reachable by the society. Each agent has its own $\mathcal{T}_A$ and T_A, but how all agents traverse the space is affected by the communication between the agents. An agent A_i communicating an artefact it has (partially) produced (or is aiming to produce) to another agent A_j may cause alterations to R_{A_j}, T_{A_j} or E_{A_j}, which in turn may result in A_j communicating back to A_i. This kind of cyclic influence can drive the exploration constantly to new areas, especially if the agents are deliberately seeking to produce artefacts seen as novel by their peers.

In a social setting such artefact exchanges between individuals may also directly affect what is created next. For example, an agent A may take an artefact produced by another agent as its own creation process' starting point (see Kantosalo and Toivonen (2016)), effectively moving to a different area in the conceptual space. This area could even be unreachable for an agent drawing inspiration only from its own productions during its creative process.

As a society's exploration of the artefact space is a conjoined process of the individual exploration processes of its agents, it is not explicitly contained in any single component of the DIFI model. Instead, it is a dynamic process where individuals aim to produce creative artefacts, but at the same time are interacting with the field which evaluates them. Further, the individuals gain information from the domain which may alter their own creative process and goals. All these properties can affect which artefacts the society can reach, i.e. $T_{\mathbf{S}}$.

$E_{\mathbf{S}}$: The Societal E for S

$E_{\mathbf{S}}$ defines which artefacts have society wide value. In some societies, $E_{\mathbf{S}}$ may have a similar relation to single agent evaluation E_A, as historical creativity (H-creativity) has to psychological creativity (P-creativity) (Boden 1992). In these societies, the artefacts evaluated highly by $E_{\mathbf{S}}$ are typically more likely to exhibit H-creativity, whereas single agent's E_A reflects merely the agent's own personal view of the artefact. However, this might not always be the case. For example, the evaluations given by a host of commoners may favour familiarity more than the evaluations given by a few experts, making it possible for $E_{\mathbf{S}}$ to become biased towards more mundane artefacts, like pop songs.

As $E_{\mathbf{S}}$ defines the artefacts which are valued by the society, it has an obvious connection to the domain in the DIFI model. The domain is a collection of cultural artefacts which have been perceived as valuable at some point during the society's lifetime. These artefacts are filtered by the field, i.e. they are subject to the social decision making policies existing in the society, and thus are also affected by the properties of the individuals.

Analysing Creative Agent Societies

The proposed extension can be utilised in various ways to describe and analyse relevant situations which may occur in creative agent societies. We begin by defining the term *substantial change* (to $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ or $E_{\mathbf{S}}$), which is utilised throughout this section. Then, we provide descriptions of the society on a single time step and how the society changes over time. We continue by showing how different agent roles, or the impacts agents have on the society, can be identified through our extension. Lastly, we sketch out some ways in which the extension can be used to analyse the policies present in the society.

We propound that the abstract analysis tools described below can be put to practice by the society's designer or analyst by defining proper (qualitative or quantitative) measures for $R_{\mathbf{S}}$, $T_{\mathbf{S}}$, $E_{\mathbf{S}}$ and substantial change to each of them. This will highlight what kind of effects particular mechanisms, policies or agent properties have on the society. We demonstrate this by presenting a practical procedure for agent role analysis in the next section. Although the formulations in this section have been done using set notation, we envision that in practice similar ideas can be formulated also using, e.g. probability calculus.

Substantial change For analysis purposes, we define the term *substantial change* (to $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ or $E_{\mathbf{S}}$), to describe an alteration which has a considerable effect on the society's operation. In CSC, this impact would preferably be of qualitative nature. For example, a substantial change to $R_{\mathbf{S}}$ could mean that the society's perception of what is considered as art would be significantly adjusted, e.g. a new painting style would be accepted as art. For $R_{\mathbf{S}}$ and $E_{\mathbf{S}}$, a necessary but not sufficient criteria for a substantial change is that some artefact which was not included before the change is included in the set after the change or vice versa. The nature of $T_{\mathbf{S}}$ is slightly different, as it considers artefacts reachable by the society. We envision that for $T_{\mathbf{S}}$ the analyst may also want to take into consideration how likely it is for the society to reach specific artefacts, as some artefacts may be reachable by several individuals.

Single Time Step Analysis

In this section, we provide some general descriptions of situations which consider the whole creative agent society $\mathbf{S}$ on a single time step t. First, we describe two situations drawn from Kantosalo and Toivonen (2016) and then two situations characterising the relation between $R_{\mathbf{S}}$ and $E_{\mathbf{S}}$.

Societal conceptual mismatch and artistic disagreement A *societal conceptual mismatch* occurs when the society $\mathbf{S}$ in its entirety cannot agree on which artefacts it considers as valid, resulting in an empty conceptual space: $R_{\mathbf{S}} = \emptyset$. Similarly, a *societal artistic disagreement* occurs when the society $\mathbf{S}$ cannot agree on which artefacts it considers as valuable, resulting in the empty set of valued artefacts $E_{\mathbf{S}} = \emptyset$.

For both of the situations, an interesting case is when $\mathbf{S}$ is composed of two separate subsocieties, which both have a non-empty internal R (or E), but are not able to form a collective understanding of it. Formally, for such societal conceptual mismatch, there exist two subsocieties $\mathbf{G}, \mathbf{H} \subset \mathbf{S}$, where $\mathbf{G} \cup \mathbf{H} = \mathbf{S}$ and $\mathbf{G} \cap \mathbf{H} = \emptyset$, such that $R_{\mathbf{G}} \neq \emptyset$, $R_{\mathbf{H}} \neq \emptyset$ and $R_{\mathbf{G} \cup \mathbf{H}} = \emptyset$.

Both societal conceptual mismatch and societal artistic disagreement may originate from multiple sources. First, the agents' individual views of R or E may be simply too different, e.g. agent A_i may not value anything even remotely close to what is valued by agent A_j, making it challenging for A_i and A_j to reconcile their views. Second, the policies present in the society may not account for minuscule changes in the agents' views, e.g. when the societal policy requires unanimity for inclusion, but every artefact is rejected by exactly one agent.

However, these situations are not always detrimental for the creative agent society. For example, in a society where the agents cultivate their expertise by specialising in particular art styles, a societal conceptual mismatch may be caused by the field being divided into specialised subfields.

Harmonious and charged societies The relation between $E_{\mathbf{S}}$ and $R_{\mathbf{S}}$ is interesting for the creative agent society as the tension between these two concepts may be a prominent reason for the society's transformation. We characterise two idealised situations which may occur in a society $\mathbf{S}$.

A society $\mathbf{S}$ is harmonious if $E_{\mathbf{S}} = R_{\mathbf{S}}$. This means that the society values exactly the artefacts it perceives as valid. This may look like a favourable situation for the society, but if the situation prolongs, it may cause stagnation as the society has reached an equilibrium where neither $E_{\mathbf{S}}$ nor $R_{\mathbf{S}}$ provokes changes in the other.

In a charged society $R_{\mathbf{S}} \neq E_{\mathbf{S}}$ and there needs to be substantial change to either $E_{\mathbf{S}}$ or $R_{\mathbf{S}}$ to make them equivalent. That is, there is some social "tension" between $E_{\mathbf{S}}$ and $R_{\mathbf{S}}$, which may provoke cultural intercourse demanding changes to either $E_{\mathbf{S}}$ or $R_{\mathbf{S}}$. This provides a natural cause for transformation in the society.

Dynamic Analysis

The temporal transformations of creative agent societies are one of the key interests in computational social creativity (Saunders and Bown 2015). These emergent, society wide phenomena can only be observed when the agent society is executed for a reasonable time span. In this section we present some fundamental characterisations for creative agent societies requiring the time dimension.

Stagnant society A society $\mathbf{S}$ has a stagnant $X_{\mathbf{S}}$, where $X_{\mathbf{S}} \in \{E_{\mathbf{S}}, R_{\mathbf{S}}, T_{\mathbf{S}}\}$, from time step t onward, if and only if for all $k > 0$ there is no substantial change from $X_{\mathbf{S}^t}$ to $X_{\mathbf{S}^{t+k}}$. That is, a stagnant society is not able to produce a substantial change to $R_{\mathbf{S}}, T_{\mathbf{S}}$ or $E_{\mathbf{S}}$ in any period of time.

For example, a stagnant $R_{\mathbf{S}}$ could mean that either the agents themselves are not able to change their own R_A, or the society's social processes cannot account for changes in the individual agent's R_A. In both cases the society wide outcome is the same, although the means to escape it differ.

A society which is stagnant in one or two of its societal elements, $E_{\mathbf{S}}$, $R_{\mathbf{S}}$ and $T_{\mathbf{S}}$, may still be able to transform in a meaningful manner. The society reaches a pathological state only when all three of these societal elements stagnate. The society ceases to evolve, limiting its ability to produce creative artefacts in the future.

Continuously transforming society A society $\mathbf{S}$ has a continuously transforming $X_{\mathbf{S}}$, where $X_{\mathbf{S}} \in \{E_{\mathbf{S}}, R_{\mathbf{S}}, T_{\mathbf{S}}\}$, if and only if for all $t \geq 0$ there exists $k > 0$ for which $X_{\mathbf{S}^{t+k}}$ is substantial change from all $X_{\mathbf{S}^a}$, where $0 \leq a \leq t$. That is, a continuously transforming society is always able to produce an alteration (to $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ or $E_{\mathbf{S}}$) which is a substantial change from all the situations (w.r.t $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ or $E_{\mathbf{S}}$) the society has previously gone through.

Diverging society The society $\mathbf{S}$ has a diverging $X_{\mathbf{S}}$, where $X_{\mathbf{S}} \in \{E_{\mathbf{S}}, R_{\mathbf{S}}, T_{\mathbf{S}}\}$, from time step t to $t+k$, if there exists two subsocieties $\mathbf{G}, \mathbf{H} \subset \mathbf{S}$, where $\mathbf{G} \cap \mathbf{H} = \emptyset$, such that $X_{\mathbf{G}^{a+1}} \cap X_{\mathbf{H}^{a+1}} \subseteq X_{\mathbf{G}^a} \cap X_{\mathbf{H}^a}$ for all $t \leq a < t+k$ and $X_{\mathbf{G}^{t+k}} \cap X_{\mathbf{H}^{t+k}}$ is a substantial change to $X_{\mathbf{G}^t} \cap X_{\mathbf{H}^t}$ and $X_{\mathbf{H}^{t+k}}$ is a substantial change to $X_{\mathbf{G}^{t+k}}$. That is, in a diverging society the two subsocieties $\mathbf{G}$ and $\mathbf{H}$ monotonously differ on their $X_{\mathbf{G}}$ and $X_{\mathbf{H}}$, ultimately producing a substantial change on what artefacts they include (in case of R and E) or can reach (in case of T). A diverging society may end up in a societal conceptual mismatch or an artistic disagreement.

Charging society The society $\mathbf{S}$ is charging from time step t to $t+k$, if $E_{\mathbf{S}^{a+1}} \cap R_{\mathbf{S}^{a+1}} \subseteq E_{\mathbf{S}^a} \cap R_{\mathbf{S}^a}$ for all $t \leq a < t+k$ holds and $E_{\mathbf{S}^{t+k}} \cap R_{\mathbf{S}^{t+k}}$ is a substantial change to $E_{\mathbf{S}^t} \cap R_{\mathbf{S}^t}$ and $E_{\mathbf{S}^{t+k}}$ is a substantial change to $R_{\mathbf{S}^{t+k}}$. That is, in a charging society $R_{\mathbf{S}}$ and $E_{\mathbf{S}}$ are diverging from each other, ultimately producing a substantial change to what artefacts they include, i.e. the society becomes charged.

Analysing Agent Roles and Impact on the Society

Analysing agent roles or their impact on the society is focal to CSC. We present two ways to describe and analyse individual agents, their relation to the whole society and their influence on it: *relation method* and *alteration method*. Both of the methods can be used either in a single time step or for a dynamic analysis of the agent society. We restrict our analysis in this section to individual agents, but it is straightforward to generalise these methods to also consider agent subsocieties.

Relation method The relation method compares the properties of a single agent A to the society wide properties, e.g. how E_A differs from $E_{\mathbf{S}}$. This method can be used to describe how the agent's properties relate to the whole society, but fails to capture the actual influence the agent has on the society via interactions and other social processes.

We define three agent types based on the single time step relation between the agent's $X_A \in \{R_A, E_A\}$ and the whole society's counterpart $X_{\mathbf{S}} \in \{R_{\mathbf{S}}, E_{\mathbf{S}}\}$: *contained*, *controversial* and *contradicting*.

A contained agent A has its own understanding of X_A fully incorporated in the society's interpretation. Formally, $X_A \subset X_{\mathbf{S}}$. That is, the society agrees on the agent's view, but the agent might not agree on other artefacts belonging to $X_{\mathbf{S}}$ (especially if $X_{\mathbf{S}}$ is a substantial change to X_A).

A controversial agent A has only a part of is own understanding of X_A incorporated in the society's interpretation. Formally, $X_A \cap X_{\mathbf{S}}$ is substantial change to X_A, $X_{\mathbf{S}}$ and $\emptyset$. This means, that there is something meaningful that is left out from the intersection from both the agent's and the society's perspective, and the intersection itself contains some meaningful set of artefacts.

A contradicting agent A has none of its own understanding of X_A incorporated in the society's interpretation. Formally, $X_A \cap X_{\mathbf{S}} = \emptyset$. In most of the cases this also implies that the agent has a conceptual mismatch (if $X = R$) or artistic disagreement (if $X = E$) with the whole society $\mathbf{S}$.

With the dynamic relation method we are able to specify if an agent A is merging or diverging from a society with respect to $X_{\mathbf{S}} \in \{R_{\mathbf{S}}, E_{\mathbf{S}}\}$. A diverging agent is actively moving away from $X_{\mathbf{S}}$ while a merging agent is actively moving inward into $X_{\mathbf{S}}$. Formally, an agent A is merging to a society $\mathbf{S}$ with respect to $X_{\mathbf{S}}$ from time step t to time step $t+k$, if it holds that $X_{A^a} \cap X_{\mathbf{S}^a} \subseteq X_{A^{a+1}} \cap X_{\mathbf{S}^{a+1}}$ for all $t \leq a < t+k$ and $X_{A^{t+k}} \cap X_{\mathbf{S}^{t+k}}$ is a substantial change to $X_{A^t} \cap X_{\mathbf{S}^t}$, and $X_{A^{t+k}}$ is a substantial change to X_{A^t}, while $X_{\mathbf{S}^{t+k}}$ is not a substantial change to $X_{\mathbf{S}^t}$. That is, X_A is actively becoming more similar to $X_{\mathbf{S}}$ while $X_{\mathbf{S}}$ does not change in a relevant manner. Ultimately, merging may cause a contradicting agent to become contained.

Alteration method The alteration method can be used to analyse what is the influence of (a property of) an agent on the whole society. In the single time step alteration method, an agent A is altered to agent A' and the society wide interpretations between the unaltered society $\mathbf{S}$ and the altered society $\mathbf{U}$ are compared. In the dynamic alteration method, the alteration from A to A' is thought to occur on a time step t, after which emergence in the unaltered society $\mathbf{S}$ and the altered society $\mathbf{U}$ are compared.

A particularly strict version of alteration method is acquired when the agent A is removed entirely from the society. We call this version of alteration method the *subtraction method*.

Sosa and Gero (2005) identify *gatekeeper* agents. These agents act as opinion leaders which have a high impact on the artefacts filtered into the domain. We formulate a gatekeeper as an agent A for which the subtraction method (or appropriately devised alteration method) produces a substantial change to $E_{\mathbf{S}}$. Formally, an agent A is a gatekeeper if $E_{\mathbf{S}\setminus\{A\}}$ is a substantial change to $E_{\mathbf{S}}$.

Sosa and Gero (2005) discuss *change agents* which are the agents driving the social change in the society. We formulate a change agent as an agent which is likely to cause a substantial change to $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ or $E_{\mathbf{S}}$, by exploiting the dynamic alteration method. Formally, an agent A is a change agent for $E_{\mathbf{S}}$ if there is substantial change from $E_{\mathbf{S}^t}$ to $E_{\mathbf{S}^{t+k}}$, but by modifying A to A' from time step t onward, thus altering the society $\mathbf{S}$ to society $\mathbf{U}$ on time step $t+1$, there is no substantial change from $E_{\mathbf{S}^t}$ to $E_{\mathbf{U}^{t+k}}$. That is, without the change agent the substantial change does not occur. (It would be beneficial to ensure that the substantial change does not occur before time step $t+k+m$ for some reasonable m, but we leave it out for clarity.)

Analysing Societal Aggregation

Our extension allows indirect observation of the effects different social policies (integrated in the societal aggregation function Π) have on the society. We sketch out a few ways in which we envision the extension to be exploited for social policy analysis.

The most straightforward way is to compare the effects two policies have on the society's $R_{\mathbf{S}}, T_{\mathbf{S}}$ and/or $E_{\mathbf{S}}$. That is, given society $\mathbf{S}$, we change some of its policies to obtain society $\mathbf{S}'$ and then compare $\{R_{\mathbf{S}}, T_{\mathbf{S}}, E_{\mathbf{S}}\}$ to $\{R_{\mathbf{S}'}, T_{\mathbf{S}'}, E_{\mathbf{S}'}\}$. This allows evaluating, e.g. if it is possible to get rid of a stagnant $R_{\mathbf{S}}$ by changing the policies within the society in a particular way.

Another way is to compare how each agent affects the whole society and to draw some insight into the policy from the comparison. For example, we can exploit a single time step subtraction method for this: for all $A_i \in \mathbf{S}$ compare $E_{\mathbf{S}\setminus\{A_i\}}$ and $E_{\mathbf{S}}$. If none of $E_{\mathbf{S}\setminus\{A_i\}}$ are a substantial change to $E_{\mathbf{S}}$, then there are no gatekeepers in the society. This could suggest that the society is somewhat balanced in its social decision making processes considering $E_{\mathbf{S}}$.

Lastly, we can compare how each agent's R, T or E relate to other agents' counterparts or to the society wide $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ or $E_{\mathbf{S}}$. For example, the socially novel goal selection (Hantula and Linkola 2018) could be characterised so that for each agent $A \in \mathbf{S}$ it is likely that for each of its peers $B \in \mathbf{S}$, where $B \neq A$, holds that E_B is a substantial change to E_A. That is, it is likely that each agent has an evaluation method which differs substantially from all other agents, or at least the agent pairs for which this does not hold are scarce.

Practical Procedure for the Alteration Method

To show that the analysis tools introduced above can be put into practice, we present an example procedure for the alteration method. The procedure allows the analyst of a creative agent society to test various hypotheses related to the agents and their effect on a specific society $\mathbf{S}$ from time step t onward, i.e. on $E_{\mathbf{S}^{t+k}}$, $R_{\mathbf{S}^{t+k}}$ and/or $T_{\mathbf{S}^{t+k}}$.

Particularly, the procedure allows the analyst to collect evidence if suspected agents (or agent properties) are indeed causing the observed emergence in the society or would the changes occur even without the agents (or agent properties) assumed to cause it. As a consequence, new hypotheses for what causes the emergence or how the dynamics of the society function may be formed.

The procedure has two main assumptions. First, it assumes that the agent society is "closed", i.e. the analyst is able to control all the parameters (and thus all the agents) within the simulation. Second, all the agents (subject to distinct hypotheses) should be alterable (even removable in the case of the subtraction method) during the simulation without breaking it.

The procedure contains the following steps:

1. Fix all random number generator seeds and other parameters of the simulation of the creative agent society $\mathbf{S}$.
2. Execute the simulation (simulation P_1).
3. Produce a hypothesis of the impact (some property of) an agent $A \in \mathbf{S}$ has on the society from a particular time step t onward.
4. Test the hypothesis by running the simulation on $\mathbf{S}$ with the same parameters again and altering A on time step t (simulation P_2).
5. Analyse how the two simulations, P_1 and P_2, differ from each other from time step t onward.
6. Repeat the steps 1-5 several times with different random number seeds to obtain statistical results.

The procedure requires the parameters to be fixed so that the exact same simulation can be run again. In this way, the only thing which changes between the two simulations is the alteration that is done to the agent. With a proper measure of (some of) the society's interpretations ($R_{\mathbf{S}}$, $T_{\mathbf{S}}$ and/or $E_{\mathbf{S}}$) it is possible to analyse if the simulations P_1 and P_2 differ substantially from time step t onward, suggesting that the alteration made to agent A was its cause. As CSC models should generally be simple (Saunders and Bown 2015), it should be reasonable to verify if this was indeed the case.

In practice, it may be challenging to devise alterations which do not change the state of the random number generators and still negate the effects of the agent properties which are under consideration. In this situation, the simulation P_2 should be run a number of times to be able to clearly conclude that the alteration produces different kind of emergence than what is observed in P_1.

Discussion

Overall, one can distinguish four different perspectives to creativity, each of which may be adopted when describing, analysing and assessing a system's creativity. The perspectives are: the creative Product (What properties make the output creative?), Process (How are creative results produced?), Producer (What kind of characteristics of systems lead to creative behaviour?), and Press (Who judges whether the product is creative, and how?) (Jordanous 2016).

Frameworks, definitions and other tools utilised in (computational) creativity typically emphasise one or two of the above mentioned perspectives. For example, the standard definition of creativity (Runco and Jaeger 2012), which defines creativity as an ability to produce outputs which are novel and valuable, highlights the Product (output) and the Producer (abilities) perspectives. On the other hand, defining computational creativity as "the art, science, philosophy and engineering of computational systems which, by taking on particular responsibilities, exhibit behaviours that unbiased observers would deem to be creative" (Colton and Wiggins 2012) puts stress on the Producer (the system taking responsibilities) and the Press (unbiased observer is the final evaluator) perspectives. Moreover, Boden's three types of creativity (Boden 1992) adopt the Process perspective by considering how the artefacts are produced.

The CSF conforms to the illustrated standard. It formalises exploratory and transformational creativity (Boden 1992), and as such it can be seen to adopt the Process perspective, taking into account also the Product perspective on a conceptual level (the validation and evaluation of the artefacts using $\mathcal{R}$ and $\mathcal{E}$). However, when it is utilised as a "stateless" conceptual framework, it is usually perceived to describe abilities of a creative system, i.e. to adopt the Producer perspective.

Following the style of the original CSF, the extension proposed in this paper is geared towards describing and analysing the overall capabilities of, or situations occurring in, a creative agent society, not the instantiated interactions taking place in the society or other transitory social phenomena. As such, the extension may be seen as counterintuitive to the CSC point of view, where observing actualised interactions between the agents during the society's execution and identifying agent attributes causing particular emergent phenomena are of key importance (Saunders and

Bown 2015). However, the society wide capabilities can be perceived as abilities of the society which *emerge* from the single agent properties (the Producer perspective) and the social policies present in the society (the Press perspective).

We propound that the extension provides a novel point of view on creative agent societies, facilitating a mathematically formal analysis of the society's creative potential and how it changes over time. Even though the extension does not offer tools to analyse the exact communications taking place in an agent society, it provides the means to analyse the consequences of communication and other social phenomena on a conceptual level. Thus, it can be used in CSC to gain evidence to show if certain mechanisms exist, which is one of the main goals of CSC models (Saunders and Bown 2015), or to test hypotheses on the impact of particular agents (or their properties) on the society. Further, by providing an alternative perspective, the extension may provide insight into new hypotheses considering the creative agent societies that would elude the analyst otherwise.

Identifying interaction emergence which does not directly affect $R_{\mathbf{S}}$, $T_{\mathbf{S}}$ or $E_{\mathbf{S}}$, such as communication patterns arising in a society of curious agents (Saunders and Gero 2002), is challenging with the extension by design. Moreover, the proposed extension is just one of the many reasonable alternatives to formulate the CSF for creative agent societies. For example, it is possible to design creative agent societies where the agents would be merely generative, but the society in its entirety would appear creative. In this paper, we have formulated each agent to contain its own CSF to show the full analysis power of the extension, but it would be possible to remove some of the CSF's elements from the agents. For example, it is common in CSC models that the agents are given by design a fixed R which does not change over time.

Conclusions

We have proposed a minimal extension to the Creative Systems Framework (Wiggins 2006a; 2006b), the CSF for creative agent societies, where each agent is an independent producer exhibiting exploratory creativity.

We have shown the extension's strength as a conceptual tool by utilising it to (1) define phenomena, both instantaneous and temporal, relevant for creativity which considers the whole agent society, (2) describe individual agent's relation to and influence over the society, and (3) characterise the effects different social policies may have on the society. All the preceding aspects can be exploited in the analysis of creative agent societies. We have demonstrated that although these analysis tools derived from the extension are conceptual, and idealised in their nature, one can devise appropriate practical procedures for them.

In the future, we hope to enhance the extension's expressive power by including more agent oriented machinery into it. For example, it would be interesting to study how each agent's memory could be presented in the extension and how it would affect the society's transformative capabilities.

Acknowledgements. This work has been supported by the Academy of Finland under grant 313973 (CACS).

References

Alvarado, J., and Wiggins, G. A. 2018. Exploring the engagement and reflection model with the creative systems frame- work. In *Proceedings of the Ninth International Conference on Computational Creativity (ICCC 2018)*, 200–207. Salamanca, Spain: Association of Computational Creativity.

Boden, M. 1992. *The Creative Mind*. London: Abacus.

Colton, S., and Wiggins, G. A. 2012. Computational creativity: The final frontier? In *Proceedings of the 20th European Conference on Artificial Intelligence*, ECAI'12, 21–26. Amsterdam, The Netherlands, The Netherlands: IOS Press.

Csikszentmihalyi, M. 1988. Society, culture, and person: A systems view of creativity. In Sternberg, R. J., ed., *The Nature of Creativity: Contemporary Psychological Perspectives*. Cambridge University Press. 325–339.

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

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

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

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

Runco, M. A., and Jaeger, G. J. 2012. The standard definition of creativity. *Creativity Research Journal* 24(1):92–96.

Saunders, R., and Bown, O. 2015. Computational social creativity. *Artificial Life* 21(3):366–378.

Saunders, R., and Gero, J. S. 2002. How to study artificial creativity. In *Proceedings of the Fourth Conference on Creativity & Cognition*, 80–87. Loughborough, UK: ACM.

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

Toivonen, H., and Gross, O. 2015. Data mining and machine learning in computational creativity. *Wiley Int. Rev. Data Min. and Knowl. Disc.* 5(6):265–275.

Wiggins, G. A. 2006a. A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems* 19(7):449–458.

Wiggins, G. A. 2006b. Searching for computational creativity. *New Generation Computing* 24(3):209–222.