Action Selection in the Creative Systems Framework

Simo Linkola[1], **Christian Guckelsberger**[2,3,4] and **Anna Kantosalo**[1,2]
[1]Department of Computer Science, University of Helsinki, Helsinki, Finland
[2]Department of Computer Science, Aalto University, Espoo, Finland
[3]Finnish Center for Artificial Intelligence (FCAI)
[4]Game AI Research Group, Queen Mary University of London, London, UK
simo.linkola@helsinki.fi, christian.guckelsberger@aalto.fi, anna.kantosalo@aalto.fi

Abstract

The Creative Systems Framework (CSF) formalises creativity as search through a space of concepts. As a formal account of Margaret Boden's descriptive hierarchy of creativity, it is at the basis of multiple studies dealing with diverse aspects of Computational Creativity (CC) systems. However, the CSF at present neither formalises action nor action selection during search, limiting its use in analysing creative processes. We extend the CSF by explicitly modelling these missing components in the search space traversal function. We furthermore integrate the distinction between a concept and an artefact, and provide stopping criteria for creative search. Our extension, the Creative Action Selection Framework (CASF), is informed by previous studies in CC and draws on concepts from Markov Decision Processes (MDPs). It allows us to describe a creative system as an agent selecting actions based on the value, validity and novelty of concepts and artefacts. The CASF brings more analytical depth for creative systems that can be modelled as utilising an action selection procedure.

Introduction

The *process* by which a creative *product* or *artefact* comes into existence represents one of the four central perspectives on creativity (Jordanous, 2016). For many Computational Creativity (CC) systems, this process can be described as ongoing selection and execution of *actions* through an *agent*. An *action selection* function hereby specifies how the agent chooses what to do next based on the current situation and its goals. Examples include a music robot's selection of musical gestures (Hoffman and Weinberg, 2010), a game character's next move (Guckelsberger, Salge, and Togelius, 2018) and a co-creative agent's choice of collaboration partners (Hantula and Linkola, 2018). Vice versa, deciding on how an agent must select its actions to be creative can mark an important step in the design of a CC system.

Despite the salience of action selection in both the analysis and design of CC systems, there has been little previous effort to include action and action selection in a specialised formal framework for creativity. Existing research on creative agents either applies heuristic action selection (e.g. Saunders, 2012; Gabora and Tseng, 2017) or considers creative action selection in more general frameworks, such as Reinforcement Learning (RL) modelled on Markov Decision Processes (MDPs) (Vigorito and Barto, 2008; Colin et al., 2016). While these approaches offer sufficient solutions for their individual purposes, they are unfitted for analysing creative processes at large: the heuristic methods miss a unified formal foundation for comparison, and generic frameworks are not sensitive to the specifics of creativity, limiting the potential for the in-depth analysis of creative processes.

In this paper, we introduce the Creative Action Selection Framework (CASF), an extension to the Creative Systems Framework (CSF) (Wiggins, 2006a,b) which can serve as the formal foundation for describing and analysing action selection in individual creative agents. It is sufficiently general to be applied to different kinds of agents, from reflex-based to learning agents (cf. Russell and Norvig, 2009).

The CSF is a formal, mathematical account of Boden's (2004) descriptive hierarchy of creativity. It defines Boden's notion of a *conceptual space* by *rules for valid concepts*. These are coupled with *rules for evaluation* applicable to concepts within and out of the conceptual space. In the centre of the search process is *a traversal function*; based on *rules for traversal*, it allows the system to move from concept to concept. The interplay between these elements can be used to define characterisations of creative search which, in turn, can be used in analysing creative systems.

The CSF however misses an account of how this search is realised through an agent's *actions*, and implicitly the *selection* of these actions. It hence has several limitations with respect to analysing creative agents: (1) it treats the traversal function as a black box and does not elaborate on how the agent decides which concept to move to next; (2) it does not distinguish between the *concept*, an agent's inner representation of an idea, and the *artefact*, the concept's external materialised expression (cf. Grace and Maher, 2015; Ventura, 2017); finally, (3) the CSF does not put forward *stopping criteria*, formalising how the agent reasons that the given concept and artefact are creative enough.

We extend the CSF to overcome these limitations, drawing on concepts in MDPs (Puterman, 2014). We (1) distinguish concepts and artefacts in the CSF. We (2) disassemble the CSF's traversal function into constituents relevant for creative agents, allowing for action selection based on the value, validity and novelty of concepts and artefacts. We (3) describe possible stopping criteria for the search of concept-artefact pairs. We reflect on the lack of universal optimality criteria for creativity, which distinguishes the CASF action selection further from optimal MDP policies.

Background and Motivation

We first introduce the CSF, identify its limitations towards analysing creative agents, and introduce MDPs as inspiration to overcoming these limitations in the CASF.

The Creative Systems Framework

The CSF (Wiggins, 2006a,b) formalises Boden's (2004) descriptive hierarchy of creativity. It hence allows for the abstract discussion of creative systems and the identification of relevant phenomena within, which particularly concerns the mechanisms of *exploratory* and *transformational creativity*. The CSF forms the basis for several studies in CC (e.g. Grace and Maher, 2015; Kantosalo and Toivonen, 2016; Alvarado and Wiggins, 2018; Linkola and Kantosalo, 2019). In this paper we focus on the exploratory part of the CSF, but the elements from our extension can also be used to drive the transformations of the system's creative behaviour.

Exploratory creativity consists of discovering novel and valuable concepts within a known conceptual space (Boden, 2004). The CSF defines it as a septuple

$$\langle \mathcal{U}, \mathcal{L}, [\![.]\!], \langle\langle ., ., . \rangle\rangle, \mathcal{R}, \mathcal{T}, \mathcal{E} \rangle, \tag{1}$$

where the individual elements of the septuple are described in Table 1. Below, we only discuss the elements which are pivotal for the rest of the paper.

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, $\top \in \mathcal{U}$.

A *function generator* $[\![.]\!]$ interprets a given rule set expressed in language $\mathcal{L}$ and outputs a function which 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 $\mathcal{E}$.

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 a mathematical theorems or buildings in a specific architecture style. $\mathcal{R}$ can be used to define the conceptual space, which Boden (2004) characterises as a structured style of thought, $\mathcal{C} = \{c \in \mathcal{U} \mid [\![\mathcal{R}]\!](c) \geq k\}$, where $k \in [0, 1]$ is a validity threshold.

The *rule set* $\mathcal{E} \subset \mathcal{L}$ defines the *evaluation* function for the system. I.e. the function generated by $[\![.]\!]$ through interpreting $\mathcal{E}$ allows to evaluate any concept $\forall c \in \mathcal{U}$ 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 output concepts, c_{out}:

$$c_{\text{out}} = \langle\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\rangle(c_{\text{in}}). \tag{2}$$

The CSF has been developed to, amongst others, describe and analyse the exploratory *capabilities* of creative systems. However, lacking the notion of actions, the CSF omits the decisions the system makes as the search unfolds. As such, it fails to describe *creative agents and their behaviour* in sufficient detail. Below, we discuss these limitations and draw on concepts from MDPs to address them.

$\mathcal{U}$	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 1: Description of the elements in the CSF

Creative Agents and the CSF

We shape our concept of *creative agents* by combining Russel and Norvig's (2009) concept of *intelligent agents* with the 'standard definition of creativity' (Runco and Jaeger, 2012). We consider a *creative agent* to be *utility-based* with the goal to produce creative, i.e. *novel* and *valuable*, *concepts* and *artefacts*. The latter distinction stems from separating an agent and its environment. A *concept* describes an agents' inner representation of ideas. An *artefact* in contrast is a materialisation of a concept as part of the agent's environment, which the agent may only have partial access to and control of[1]. The same concept can be expressed as different artefacts using diverse skills or means: the concept of a flying horse can be expressed as a poem or as a painting, and in both domains there are a plethora of ways to do so[2].

During a creative agent's operation, the system's overall state can be conceived as a tuple of two states: one for the concept and one for the artefact, both of which can be "empty" at any given time. We denote this as the agent's *position*. To produce concepts and artefacts, the agent performs *actions* that affect the state of its concept, the artefact's, or both. The search for concepts happens in the concept state space, whereas artefacts are manipulated in the artefact state space[3]. The effectiveness of actions to induce change depends on the dynamics of these spaces.

We define the agent's overall goal as finding a position where both the present concept and the artefact are assessed favourably, and the artefact is a good fit for the concept, i.e.

[1]This environment can be arbitrarily complex. Here, we only consider the part constituting the agent's artefact.

[2]Our concept-artefact distinction is informed by Grace and Maher (2015) and Ventura's (2017) genotype-phenotype distinction.

[3]We acknowledge that this distinction of an internal conceptual space and an external environment with artefacts is a simplification as from a *monist* position (Jaworski, 2011), there is no difference between these domains in terms of substance. Moreover, the existence of a clear boundary separating an agent and its external environment is disputed, challenging the very concept of agency.

it expresses the concept well. Crucially, the agent may first thrive to find a prominent concept and then express it as an artefact. Alternatively, it may seek an artefact which is then associated with a concept, as in the case of Duchamp's "ready-mades" where an existing artefact is placed in a new context in which it can be conceived in a fundamentally different way. Moreover, the agent may alternate between concept and artefact exploration, incorporating insights from perceiving the unfinished artefact along the way.

In existing CC systems, the exploration of potential concepts and artefacts is often interrupted externally by the user or designer as soon as they are satisfied by the outcome. We however model a creative agent's own perspective on its concept and artefacts. Hence, we must formalise how the agent itself decides that it should stop its creative process.

We identify several shortcomings of the CSF towards describing *creative agents* as introduced above:

1. *No actions:* The agent's actions are obscured in the traversal function and the CSF lacks a formal account of the agent's action selection in the creative process.
2. *No concept-artefact separation:* The CSF does not distinguish between concepts and artefacts and hence cannot explicate how their possible relationships can impact the agent's creative process.
3. *No stopping criteria:* The CSF offers little explanation of *when* a creative agent should "stop" its exploration, e.g., to start anew or to output its current concept-artefact pair.

We address these limitations by drawing inspiration from the framework of Markov Decision Processes.

Markov Decision Processes

A Markov Decision Process (MDP) (Bellman, 1957) describes a *sequential decision-making problem*. It models the possible interaction between an arbitrary agent and its environment over time, where the environment is distinguished from the agent by everything that they cannot change arbitrarily (Sutton and Barto, 2018, p. 50). At each point of the interaction, the agent can receive a reward from the environment. The problem consists of finding a *policy*, i.e. a decision-making rule, which maximises cumulative future reward from an initial state onward (Puterman, 2014, p. 2).

A (infinite horizon) MDP is a quadruple $\langle \mathcal{S}, \mathcal{A}, p, \rho \rangle$ (Puterman, 2014, pp. 1-2). The environment state at time-step $t{\geq}0$ is denoted as $s^t{\in}\mathcal{S}$. An agent's action $a^t{\in}\mathcal{A}$ can influence the future state of the environment s^{t+1} determined by *environment dynamics* $p : \mathcal{S}{\times}\mathcal{A}{\times}\mathcal{S}{\to}[0,1]$ given as a conditional probability distribution $p(s^{t+1}|s^t, a^t)$, where the actions available to an agent can depend on the current environment state. The actions produce an immediate reward signal $r^{t+1}{\in}\mathbb{R}$ determined by the *reward function* $\rho : \mathcal{S}{\times}\mathcal{A}{\times}\mathcal{S}{\to}\mathbb{R}$, given by $\rho(s^{t+1}, a^t, s^t){=}r^{t+1}$. The Markov assumption implies that s^t must encode all relevant information about the past agent-environment interaction that matters for the future dynamics.

All previous elements formalise the decision-making *problem*, but do not provide a *solution*. Specifying the probability of the agent selecting a specific action in a specific state $\pi(a|s)$, a *policy* is an attempt to solve the problem specified by a MDP. A solution to the problem is an *optimal policy* π^* that maximises cumulative future reward. Crucially, there can be several optimal policies.

MDPs and methods to (approximately) solve them, e.g. RL, have been previously discussed and utilised in creative contexts (e.g. Hoffman and Weinberg, 2010; Hantula and Linkola, 2018). Vigorito and Barto (2008) model creativity as a blind variation-and-selection process, arguing that creative behaviour may be acquired by hierarchical RL as means to reduce the complexity of creative search. Colin et al. (2016) present one possible mapping of Ritchie's (2012) "simplified version" of the CSF to hierarchical RL where the MDP policies are CSF's concepts and their evaluation in the CSF is likened to the discounted return of the policy. Based on this mapping, they argue that hierarchical RL realises creative behaviour in the form of exploratory and transformational creativity.

We next move beyond existing work by proposing an extension to the CSF which accounts for actions and action selection by adopting the notions of *states*, *actions* and *environment dynamics* from MDPs. We also elaborate on the use of these notions to fit the specifics of creative agents, in particular the distinction between concepts and artefacts.

Bringing Actions to the CSF

We now introduce the Creative Action Selection Framework (CASF) as extension to the CSF to leverage its power for the description of creative agents. We consider a *single* creative agent with a *closed loop* creative process. That is, the agent moves in the search space using solely its own actions and reasoning. The agent's actions can change the states of the concept and artefact spaces, which the agent can then observe and assess. We thus do not explicitly consider any *co-creative* and *interactive agents* which rely on external feedback or communication with other agents. However, most of our formulation also applies to these cases.

Our main contribution in this section is the deconstruction of the agent's traversal function $\langle\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\rangle$. Before we can address this however, we must add detail to some of the CSF's other elements: we introduce the concept of *actions* into the traversal rules $\mathcal{T}$, and make *time* and the *concept-artefact distinction* explicit in the framework.

Universe $\mathcal{U}$: We distinguish two subsets from the universe: $\mathcal{U}_\omega$, which encompasses concepts, and $\mathcal{U}_\alpha$, which contains artefacts. $\mathcal{U}_\omega$ and $\mathcal{U}_\alpha$ comprise the possible states of concept and artefact search, respectively. The empty state $\top$ is a member of both subsets.

Input and output sequences c_{in} and c_{out}: As we view a creative agent in a combination position of a concept and an artefact, we consider each element in the input and output sequences c_{in} and c_{out} as a position tuple $\phi = (\omega, \alpha)$, where $\omega \in \mathcal{U}_\omega$ and $\alpha \in \mathcal{U}_\alpha$. Thus, for a single element sequence, an agent's traversal (Equation 2) may be denoted by

$$\phi_{\text{out}} = \langle\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\rangle(\phi_{\text{in}}). \quad (3)$$

In the rest of this paper, we mostly consider such singular input and output positions. However, as $\mathcal{U}$ can represent *anything*, it can also comprise *sequences* of such positions. Our formalisation works mostly for both cases.

Time $t \geq 0$: Time is implicitly present in the original CSF's ordering of input and output sequences. We make it explicit in an agent's closed loop creative process by denoting $\phi_{\text{in}} = \phi^t$ and $\phi_{\text{out}} = \phi^{t+1}$. That is, we assume that time progresses with traversal from t to $t+1$, and that the output of the last time step's traversal function is fed as an input to the traversal function in the next time step.

Value, validity and novelty: Novelty is not explicitly considered in the original CSF – what the evaluation rules $\mathcal{E}$ entail is left vague. To make novelty explicit, we denote the rules for novelty by $\mathcal{E}_N$[4], and rules for evaluation by $\mathcal{E}_E$. Furthermore, we modify both the interpreted validity and evaluation function ($[\![\mathcal{R}]\!]$ and $[\![\mathcal{E}_E]\!]$) to accept a *position* as input and address novelty by adding an interpreted novelty function $[\![\mathcal{E}_N]\!]$ which also operates on a position. We assume that these functions return a triplet of real values representing the assessment of the input concept, the artefact and their combination. This results in the following functions:

$$\text{evaluate}(\phi) = [\![\mathcal{E}_E]\!](\phi) = (e_\omega, e_\alpha, e_{\omega\alpha}) \tag{4}$$
$$\text{validate}(\phi) = [\![\mathcal{R}]\!](\phi) = (v_\omega, v_\alpha, v_{\omega\alpha}), \text{ and} \tag{5}$$
$$\text{novelty}(\phi) = [\![\mathcal{E}_N]\!](\phi) = (n_\omega, n_\alpha, n_{\omega\alpha}), \tag{6}$$

where the subscript ω denotes assessment for the concept, the subscript α assessment for the artefact, and the subscript $\omega\alpha$ assessment for the combination[5].

The combination assessments describe if the artefact expresses the concept properly (validity) and how elegant the expression is (evaluation). For example, for the concept of *freedom*, valid artefacts could portray unlocked shackles or a bird. However, if the bird happened to be a penguin, the concept might not be as elegantly expressed as using a more stereotypical bird in a proper context.

For novelty, as ϕ is a single position for which novelty is computed, we assume that the function incorporates the selected input positions (such as those that the agent chooses to output) into a persistent, internal model of novelty for successive calls. That is, the novelty is computed with respect to the history of the agent. Novelty of the combination serves here a similar purpose as the combination assessments in validity and evaluation functions: it may be used to assess how new the concept-artefact pair is.

Traversal rules $\mathcal{T}$: Traversal rules govern how the system explores the search space. As we model the agent's exploration as an action selection process inspired by MDPs, we distinguish the following subsets in $\mathcal{T}$:

$\mathcal{T}_\mathcal{A}$ (or simply $\mathcal{A}$): the actions available to the agent;

$\mathcal{T}_\phi$: the agent's model of the search space (how the concept and artefact spaces react to the agent's actions); and

$\mathcal{T}_\pi$: the policy specifying the habit to move in both spaces.

[4]Specific traversal rules can foster novelty, but it is ultimately the result of evaluation. In contrast to Grace and Maher (2015), we hence only consider it in evaluation and not in the traversal rules.

[5]Further distinctions are possible, e.g. the evaluation of the artefact given the concept, $e_{\alpha|\omega}$, and vice versa, $e_{\omega|\alpha}$.

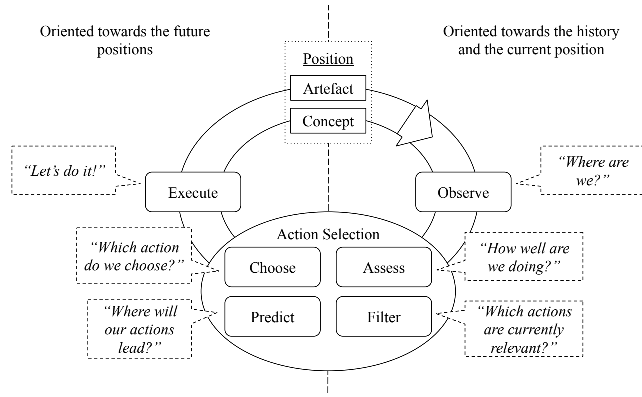

Figure 1: Closed loop traversal: an agent (1) observes its position, (2) selects the next action and (3) executes it, which may influence the position. Different action selection mechanisms use some or all of the following components in varying order: the assessment of the present position, the filtering of actions into a relevant subset, the prediction of action outcomes, and the choosing of an action.

The actions $a \in \mathcal{A}$ form the nucleus of action selection and of our deconstruction of the traversal function. An action may alter the agent's place in the concept space, in the artefact space, or in both. By construction, CASF also includes two special action types present in creative processes and previously discriminated in CC, *translation* and *re-perception* (Grace and Maher, 2015; Ventura, 2017). The agent is completely free to execute which action it chooses, but the actions available to the agent may differ from position to position.

Deconstructing the Traversal Function

Next, we go through a single iteration of the agent's closed loop traversal encompassed in the traversal function $\langle\!\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\!\rangle$, and illustrated in Figure 1. The agent initially *observes* its current position, and eventually *executes* the next action. The action selection can be implemented via different mechanisms; each uses some or all of the following components, potentially in different order: the agent *assessing* the position's value, validity and novelty, *filtering* its own action possibilities, *predicting* their possible outcomes, and *choosing* the next action to take. Below, we address these six components individually and separate them into their own functions. This provides us with conceptual clarity, and supports the CASF's use in the description and analysis of diverse creative agent types, including hard-coded reflex-based agents and those capable of learning from experience to adapt action selection. We acknowledge that this separation does not account for all atomic elements in creative action selection – e.g. *predicting* could be further divided into predicting the next positions and assessing them.

Observe (position): The agent needs to observe its position at time t to gather information about where it is in the concept and artefact spaces. Formally,

$$\phi_{\text{obs}} := \text{observe}(\phi^t), \tag{7}$$

where ϕ^t is the agent's actual position at time t and $\phi_{\text{obs}} =$

(ω^t, α^t), $\omega^t \in \mathcal{U}_\omega$ and $\alpha^t \in \mathcal{U}_\alpha$, denotes its observation by the agent. Generally, this observation is *imperfect*, i.e. $\phi^t \neq \phi_{\text{obs}}$; the agent may not have a full view of the universe, and observe some of the dimensions of $\mathcal{U}$ (relevant for either concepts or artefacts or both) incorrectly or not at all. In the rest of this section, we typically assume that ϕ_{obs} is a single position. However, whenever we deal with mechanisms which include the agent's history, we assume that ϕ_{obs} contains the agent's (recent) observed position history or the agent has other means to retrieve it from data structures which are not explicitly specified.

Assess (position): The agent may assess its observed position. We denote this assessment by

$$\delta := \text{assess}(\phi_{\text{obs}}, \mathcal{R}, \mathcal{E}) = \begin{bmatrix} [\![\mathcal{E}_E]\!](\phi_{\text{obs}}) \\ [\![\mathcal{R}]\!](\phi_{\text{obs}}) \\ [\![\mathcal{E}_N]\!](\phi_{\text{obs}}) \end{bmatrix}^T, \tag{8}$$

where the three functions on the right are defined in Equations 4-6. These assessments are essential to direct the agent's further reasoning process. For example, if all assessments are favourable, the agent may choose to output the concept-artefact pair; if only the concept is assessed favourably, the agent may only develop the artefact further.

Filter (actions): The agent may filter which actions are appropriate given its current position and the assessments:

$$\mathcal{A}_{\text{appr}} := \text{filter}(\phi_{\text{obs}}, \Delta, \mathcal{T}_{\mathcal{A}}), \tag{9}$$

where $\mathcal{A}_{\text{appr}} \subset \mathcal{T}_{\mathcal{A}}$ is the set of actions the agent finds appropriate with respect to its goal(s), and Δ is either the returned assessments, δ, or (parts of) the assess-function itself.

The filtering step adheres to the "hard traversal rules" of the system which can not be broken in any circumstance, but it may also be based on the agent's (recent) history and goals. Typical cases of the former are the hard-coded restrictions for the system to stay in the conceptual space, e.g., by restricting the percentage of the canvas that can be red. The latter part can serve similar purposes as "focus" or "attention" in animals. If the agent's current goal e.g. is to compose a painting using only triangles, then the appropriate actions in the artefact space may deal only with triangles.

Predict (actions): The agent may inform its next action by predicting the consequences of available, potentially filtered actions. This may involve a (learned) model of action outcomes, or some fixed heuristics. Below, we assume the former case, where the agent uses its model of the exploration dynamics, $\mathcal{T}_\phi$, an adaptation of the environment dynamics from MDPs. Formally, these dynamics $\mathcal{T}_\phi : \mathcal{U}_\omega \times \mathcal{U}_\alpha \times \mathcal{A} \times \mathcal{U}_\omega \times \mathcal{U}_\alpha \rightarrow [0, 1]$ are given by a conditional probability distribution $\mathcal{T}_\phi(\omega^{t+1}, \alpha^{t+1} | \omega^t, \alpha^t, a^t)$. The predict function for an arbitrary action is given by

$$(\hat{\phi}^{t+1}, \hat{e}^{t+1}, \hat{v}^{t+1}, \hat{n}^{t+1}) := \text{predict}(\phi_{\text{obs}}, a, \mathcal{R}, \mathcal{E}, \mathcal{T}_\phi), \tag{10}$$

where $\hat{\phi}^{t+1}, \hat{e}^{t+1}, \hat{v}^{t+1}$, and $\hat{n}^{t+1}$ are the predicted observed position in the next time step caused by the action and its predicted value, validity and novelty triplets, respectively. In general, the output[6] can also be a sequence of $(\hat{\phi}^{t+1}, \hat{e}^{t+1}, \hat{v}^{t+1}, \hat{n}^{t+1}, \hat{p})$ tuples, where $\hat{p}$ is the predicted probability for that outcome. We hat these variables to indicate that they represent potential rather than actual states.

Where the filter-function restricted the appropriate action set with respect to the agent's history and goals, the predict-function envisions likely outcomes of the actions and approximates their assessments. The predictions may be inaccurate and incomplete as the agent's model of the exploration dynamics may only encode some of the relevant concept and artefact dimensions well. Strong predictive capabilities allow the agent to assess an action more thoroughly by exploring its consequences further into the future. Informally, an agent employing a predictive model "imagines" possible concepts and artefacts before they are realised, and can use that imagination to drive its creative process.

Apt prediction capabilities are essential when *executing* an action (see below) is resource demanding or non-reversible. For example, a painting robot may take a considerable time in executing a set of instructions to paint the next patch and it costs money to buy the paint needed. Moreover, in safety critical domains, the actions may cause dangerous situations and ultimately harm humans or other actors.

Choose (action): The last part of the action selection is naturally about choosing the next action to take. This may be entirely random, but can also be informed by the current position's *assessment* and by *predictions* of potentially *filtered* actions. Moreover, it could take into account $\mathcal{T}_\pi$, (learned) heuristics on how to proceed from this (or similar) observations onwards as adaptation of the MDP *policy*[7]. As a creative agent is constantly in a combination of two states, we have $\mathcal{T}_\pi : \mathcal{U}_\omega \times \mathcal{U}_\alpha \times \mathcal{A} \rightarrow [0, 1]$, which is given as a conditional probability distribution $\mathcal{T}_\pi(a^t | \omega^t, \alpha^t)$[8].

The choose-function incorporates these information sources into a mapping to a single chosen action. Formally,

$$a := \text{choose}(\phi_{\text{obs}}, \Delta, \mathcal{D}_{\text{preds}}, \mathcal{T}_\pi), \tag{11}$$

where Δ denotes either the returned assessments, δ, or (parts of) the assess-function itself, and $\mathcal{D}_{\text{preds}}$ is a potential mapping from actions to their predictions obtained using the predict-function. If the agent can only filter, the mapping contains only the keys for each $a \in \mathcal{A}_{\text{appr}}$.

Creative agents can implement the choose-function in different ways. It may use well known methods, e.g. a *softmax* (Sutton and Barto, 2018, p. 322) over action assessments, or it can incorporate precoded heuristics. Moreover, the function may depend on the recent history of the agent, e.g., to determine the priority of novelty, value and validity, or to compare the assessments of the current position and the predicted outcomes with past assessments.

Execute (action): Executing the chosen action is the agent's means to potentially impact its position. Formally,

$$\phi^{t+1} := \text{execute}(a), \tag{12}$$

where ϕ^{t+1} is the objective position at the next time step. This is the output of the traversal function.

[6] The predict-function may also return confidences for the predictions, which are argued to be required for assessing surprise (Grace and Maher, 2015) of the action outcomes.

[7] Considering the policy as argument to *choose* allows us to consider *off-policy* traversal and policy updates in the CASF.

[8] This formalisation also accounts for deterministic policies by modelling them as Dirac delta distributions.

Algorithm 1 Example of model-free action selection in computing the traversal function $\phi^{t+1} = \langle\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\rangle(\phi^t)$.

$\phi_{\text{obs}} \leftarrow \text{observe}(\phi^t)$
$\delta \leftarrow \text{assess}(\phi_{\text{obs}}, \mathcal{R}, \mathcal{E})$
store (a^{t-1}, δ) in dictionary $\mathcal{D}_\delta$
$a^t \leftarrow \text{choose}(., ., \mathcal{D}_\delta, .)$
$\phi^{t+1} \leftarrow \text{execute}(a^t)$

Algorithm 2 Example of model-based action-selection in computing the traversal function $\phi^{t+1} = \langle\langle \mathcal{T}, \mathcal{R}, \mathcal{E} \rangle\rangle(\phi^t)$.

$\phi_{\text{obs}} \leftarrow \text{observe}(\phi^t)$
$\mathcal{A}_{\text{appr}} \leftarrow \text{filter}(\phi_{\text{obs}}, ., \mathcal{T}_{\mathcal{A}})$
foreach $a \in \mathcal{A}_{\text{appr}}$ **do**
  $(\hat{\phi}^{t+1}, \hat{e}^{t+1}, \hat{v}^{t+1}, \hat{n}^{t+1}) \leftarrow \text{predict}(\phi_{\text{obs}}, a, \mathcal{R}, \mathcal{E}, \mathcal{T}_\phi)$
  store $(\hat{\phi}^{t+1}, \hat{e}^{t+1}, \hat{v}^{t+1}, \hat{n}^{t+1})$ in dictionary $\mathcal{D}_{\text{preds}}$
end
$a^t \leftarrow \text{choose}(\phi_{\text{obs}}, ., \mathcal{D}_{\text{preds}}, \mathcal{T}_\pi)$
$\phi^{t+1} \leftarrow \text{execute}(a^t)$

The next traversal cycle starts with the agent observing its potentially new position as basis for action selection and to verify whether the previous action has resulted in a desirable or predicted outcome. This is not certain: the agent's predictions may be defective, the dynamics of the performed action may be noisy, e.g. when splattering paint, or the execution may add imperfections to the planned action, e.g. by an unintentional rotation of a loose joint.

There exist many means to combine the above building blocks into a specific action selection function: many creative agents do not use all elements, or their functionality overlaps in the specific implementation. To illustrate the spectrum of possible approaches and support the applicability of the CASF across different agent types, we provide pseudocode for two action selection mechanisms in the traversal function. In the *model-free* approach in Algorithm 1, the agent chooses the next action to perform based on a past record of action-assessment tuples. A concrete example of this approach is *Q-learning* (Sutton and Barto, 2018, p. 131 ff.). In the *model-based* approach in Algorithm 2 in contrast, the agent predicts the consequences of each action in a set of previously filtered, appropriate actions. It then leverages these future action assessments to choose the next action to execute. An example for this is *Monte Carlo Tree Search* (Sutton and Barto, 2018, p. 185 ff.).

Stopping Exploration

By repeatedly invoking its traversal function based on the previous cycle's output, the agent moves in the concept and artefact space. Crucially, the original CSF does not specify any *stopping criteria* for exploratory creativity to explain how the agent reasons that it has arrived at a particularly apt concept-artefact pair, which it could, for example, then show to others. Below, we describe a few potential stopping criteria, partly informed by the additional elements in the CASF:

Thresholds: A simple way for the agent to decide that the given concept-artefact pair is creative enough is that all the assessments are above some *absolute thresholds* given to the agent during its initialisation. This relates to Wiggins's (2006a) usage of a filtering thresholds for the conceptual space and the set of valued artefacts. *Dynamic thresholds* work in a similar manner, but instead of the thresholds being fixed, the agent may alter them based on its own history and experience on assessing concepts, artefacts and their combinations. This gives the agent more room e.g. to determine acceptable assessments in a certain area of the search space.

Predictions: Given a predictive model, the agent may approximate if it could, by means of its acting, cause a concept-artefact pair in the near future that is assessed more favourably than the current one. If the likelihood for this is high, then the agent may reason that it should not stop in the current position. If the likelihood for generating a better position is low however, the agent may either output the position (if its overall assessment is favourable) or start anew (if the current position is below average). Predictive reasoning and stopping is especially important in creative domains where actions can hardly be reversed, e.g. when painting on a physical canvas, or in music and dance improvisation.

Resource restrictions: The agent may stop exploration based on the consumption of a specific, tracked resource. For instance, it may have a time budget, and stop as soon as additional exploration is not predicted to yield any improvements in assessing the current concept-artefact pair.

Other criteria: Naturally, there are a multitude of other possible stopping criteria for the agent based on its design and purpose. Moreover, the criteria above may be combined.

Action Selection in Analysis

Next, we consider how the CASF may be *applied* in describing and analysing creative agents. Due to the scope of the paper, we refrain from most mathematical formulations and merely list dimensions which may provide a useful starting point for an in-depth analysis of an agent's creative process.

We first distinguish dimensions that enable, amongst others, the coarse analysis of a creative agent's action capabilities: movement in the position space (Do the actions cover movement in both the concept and the artefact space?), action possibilities (How many actions are available to the agent in any given situation? How many of these result in distinct outcomes?), action granularity (How fine-grained are the agent's potential movements in the concept and artefact spaces?), execution control (Does executing an action in a certain position map to one or to multiple outcomes?), action scope (Does the agent have actions that are fundamentally different, e.g. actions for producing both music and visual art?), and action learning (Is the agent able to learn new or more complex actions during its creative process?).

The dimensions above may serve as the basis of analysing a creative agent. However, they neither provide sufficient detail on the characteristics of the action selection procedure inside the traversal function, nor do they govern the agent's overall process of moving in the position space (i.e. action-position sequences) while searching for apt concept-artefact pairs. We address (parts of) both of these cases below.

Filtering characteristics: Does the set of appropriate actions change between positions? How strict is the agent's filtering? Too strict filtering may restrict creativity, while lenient filtering may hinder performance. Controlled oscillation between lenient and strict filtering may hint that the agent's creative process can be characterised with cycles of divergence and convergence.

Prediction abilities: Is the agent able to predict outcomes of its actions? How well the agent's prediction matches the outcome of the executed action? How reliable are the predicted value, validity and novelty? Prediction ability is important for the agent to work towards its goals, while mistakes in predictions may give arise to serendipity.

Flexibility of the process: Does the system's operation result in similar action sequences or does the action sequences differ between producing multiple concept-artefact pairs for output. Is the process largely the same no matter which concept and/or artefact is produced? Overall flexibility of the process can be seen as desirable, but it should be coupled with the reliability of the process.

Reliability of the process: How reliably the creative agent produces apt concept-artefact pairs? How well the agent is able to exploit promising subspaces of the whole position space? This dimension analyses the agent's alignment towards its overall goal. It should be used in conjunction with the other dimensions to ensure that the agent does not simply wander around.

Lastly, we reflect on the deterrent modes of creativity which deal with the system's overall deficiencies to reach apt concepts (Wiggins, 2006a). With the CASF, we can describe these modes from the agent's perspective. An agent which only momentarily visits any of these deterrent modes can be argued to be aligned towards its goal of producing apt concept-artefact pairs. However, some exploration in these modes may be needed in order for the agent to reach apt positions which would not be discovered otherwise.

In order for the original mode descriptions to be applicable for creative agents, we need to modify them (1) to take time into account and (2) to allow concept-artefact pairs. We use *generative uninspiration* (Wiggins, 2006a) as an example, but similar modifications can be done to all of the modes. For simplicity, we only cover the case of evaluating the concept-artefact combination, but any composition of concept and artefact assessments could be considered. Below, we use *time leniency*, denoted by m, to mark the agent's own understanding of what is an acceptable number of time steps to continuously spend in a deterrent mode.

Definition. *Generative Uninspiration* Let ϕ^t be an agent's observation of its position in the time step t and let $[\![\mathcal{E}_E]\!](\phi^t) = (e^t_\omega, e^t_\alpha, e^t_{\omega\alpha})$ be the agent's evaluations for its observed position in the time step t. An agent with time leniency m notices itself exhibiting generative uninspiration (for concept-artefact pairs) on time step $k \geq t + m$, if $\forall e_{\omega\alpha} \in (e^t_{\omega\alpha}, \ldots, e^k_{\omega\alpha}) : e_{\omega\alpha} < \epsilon$, where ϵ is the agent's threshold for valuable concept-artefact combinations.

Discussion

We have defined a creative agent's main goal as the production of creative, i.e. novel and valuable, concept-artefact pairs. We have moreover formalised the respective process as ongoing selection and execution of actions based on the assessment of these pairs, resulting in a sequence of action-position tuples. The CASF hence provides a deeper account of the *product* and *process* perspective on creativity (Jordanous, 2016) than the original CSF, in that it characterises the desired properties of creative products, and uses them to ground the creative process in action-selection. Our extension accounts for the *person* perspective by considering the extent of the agent's creative capabilities in a similar but more detailed manner as the original CSF's description of deterrent modes of creativity. For instance, the CASF allows us to ask if the agent has the *ability* to predict the consequences of its actions.

The concept-artefact distinction in the CASF is included for completeness, as it is present in theories of human creative processes and has been discriminated before in CC (Grace and Maher, 2015; Ventura, 2017). However, not all agents are capable of using both state spaces. In these cases, the framework can be reduced to deal only with concepts or artefacts.

Creative processes can also be considered through the FACE model (Colton, Charnley, and Pease, 2011), which aims to describe progress in the development of a creative system. The FACE model distinguishes between *concepts* (a piece of code, e.g. a function) and *expressions* generated using these concepts. This bears some similarity to *concepts* in the CASF and their expressions as *artefacts*. However, the FACE model handles the expression generation as a single act (similar to Ventura, 2017), whereas in the CASF an artefact may be generated using multiple actions. Moreover, a CASF agent may search for an apt concept for a fixed artefact, a case that is not accounted for in the FACE model.

In the CASF, the observe-function is treated as a black box. However, certain theories of creative processes, such as *engagement-reflection* (Sharples, 1996), treat the perception of an artefact as an action in its own right. While the CASF encompasses re-perception actions that describe an artefact as a concept, it does not explicate all the different types of perceiving actions, some of which may involve assessing the artefact in a particular way. We believe that the framework would benefit from making them explicit.

Surprise is a prominent subjective assessment which is sometimes considered a defining characteristic of creativity (Boden, 2004), and which may influence action-selection and hence the direction of a creative process. We have not considered surprise in this paper, but it could be included into the CASF by e.g. adopting Grace and Maher's (2015) proposal: the predict-function can be modified to return confidence values for the predictions and the prediction accuracy can then be determined in the next time step. If the prediction confidence was high and the observed position or an assessment in the next time step deviates sufficiently from the prediction, the agent may quantify this as surprise.

Conclusions

We have extended the CSF to distinguish concepts and artefacts and to include action selection in its traversal function. The resulting Creative Action Selection Framework (CASF) is the first formal framework oriented towards the analysis of creative *agents*. It provides formal tools to analyse a creative agent's on-going *process* of producing concepts and artefacts by dividing the agent's traversal of the search space into six conceptually separate building blocks which may each be further specified and analysed. By formalising the agent's action and position sequences and their alignment with their goals, the CASF moreover also affords a more high-level analysis of the resulting behaviour.

One goal for future work is to detail the formalism further so that the components in an agent's action selection can be considered in terms of simpler, shared atomic factors. This would highlight their commonalities and support the analysis and comparison of creative systems. By introducing action-selection to the CSF, the CASF allows us to compare a dedicated framework for the analysis of creative agents to more general decision-making frameworks in AI. Another goal for future work is to investigate such possible mappings and consequently evaluating the specific notions of creativity developed by Boden (2004) and formalised by Wiggins (2006a,b) in a wider AI context.

Acknowledgments

SL is funded by Academy of Finland grants 313973 (CACS), and 328729 (CACDAR). AK is funded by Academy of Finland grant 311090 (Digital Aura) and was funded by by the doctoral programme in Computer Science at the University of Helsinki. CG is funded by the Academy of Finland Flagship programme "Finnish Center for Artificial Intelligence" (FCAI) and was funded by EPSRC grant [EP/L015846/1] (IGGI).

References

Alvarado, J., and Wiggins, G. A. 2018. Exploring the engagement and reflection model with the creative systems framework. In *Proc. 9th Int. Conf. Computational Creativity*, 200–207.

Bellman, R. 1957. A Markovian decision process. *Journal of Mathematics and Mechanics* 679–684.

Boden, M. A. 2004. *The creative mind: Myths and mechanisms*. Routledge.

Colin, T. R.; Belpaeme, T.; Cangelosi, A.; and Hemion, N. 2016. Hierarchical reinforcement learning as creative problem solving. *Robotics and Autonomous Systems* 86:196–206.

Colton, S.; Charnley, J.; and Pease, A. 2011. Computational Creativity Theory: The FACE and IDEA Descriptive Models. In *Proc. 2nd Int. Conf. Computational Creativity*, 90–95.

Gabora, L., and Tseng, S. 2017. The social benefits of balancing creativity and imitation: Evidence from an agent-based model. *Psychology of Aesthetics, Creativity, and the Arts* 11(4):403–419.

Grace, K., and Maher, M. L. 2015. Specific curiosity as a cause and consequence of transformational creativity. In *Proc. 6th Int. Conf. Computational Creativity*, 260–267.

Guckelsberger, C.; Salge, C.; and Togelius, J. 2018. New And Surprising Ways to be Mean: Adversarial NPCs with Coupled Empowerment Minimisation. In *Proc. Conf. Computational Intelligence in Games*. IEEE.

Hantula, O., and Linkola, S. 2018. Towards goal-aware collaboration in artistic agent societies. In *Proc. 9th Int. Conf. Computational Creativity*, 136–143.

Hoffman, G., and Weinberg, G. 2010. Gesture-based human-robot jazz improvisation. In *Proc. Int. Conf. Robotics and Automation*, 582–587. IEEE.

Jaworski, W. 2011. Mind-body theories and mind-body problems. In *Philosophy of Mind: A Comprehensive Introduction*. Wiley-Blackwell.

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 *Proc. 7th Int Conf. Computational Creativity*, 77–84.

Linkola, S., and Kantosalo, A. 2019. Extending the creative systems framework for the analysis of creative agent societies. In *Proc. 10th Int. Conf. Computational Creativity*, 204–211.

Puterman, M. L. 2014. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley.

Ritchie, G. D. 2012. A closer look at creativity as search. In *Proc. 3rd Int. Conf. Computational Creativity*.

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

Russell, S., and Norvig, P. 2009. *Artificial Intelligence: A Modern Approach*. USA: Prentice Hall Press, 3rd edition.

Saunders, R. 2012. Towards autonomous creative systems: A computational approach. *Cognitive Computation* 4(3):216–225.

Sharples, M. 1996. An account of writing as creative design. *The science of writing* 127–148.

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

Ventura, D. 2017. How to Build a CC System. In *Proc. 8th Int. Conf. Computational Creativity*, 253–260.

Vigorito, C. M., and Barto, A. G. 2008. Hierarchical representations of behavior for efficient creative search. In *AAAI Spring Symposium: Creative Intelligent Systems*, 135–141.

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.