Chapter 7
An Enactive Model of Creativity for Computational Collaboration and Co-creation

Nicholas Davis, Chih-Pin Hsiao, Yanna Popova, and Brian Magerko

7.1   Introduction

The modern landscape of computing has rapidly evolved with breakthroughs in new input modalities and interaction designs, but the fundamental model of humans giving commands to computers is still largely dominant. A small but growing number of projects in the computational creativity field are beginning to study and build creative computers that are able to collaborate with human users as partners by simulating, to various degrees, the collaboration that naturally occurs between humans in creative domains (Biles 2003; Lubart 2005; Hoffman and Weinberg 2010; Zook et al. 2011; Davis et al. 2014). If this endeavor proves successful, the implications for HCI and the field of computing in general could be significant. Creative computers could understand and work alongside humans in a new hybrid form of human-computer co-creativity that could inspire, motivate, and perhaps even teach creativity to human users through collaboration.

To reach this optimistic future, the field of computational creativity needs a conceptual framework and model of creativity that can account for the collaborative and improvisational nature of human creativity. Traditional cognitive science theories view cognition and creativity as an abstracted manipulation of symbols that happens solely in the brain (e.g., Newell et al. 1959). The new cognitive science

N. Davis (✉) • B. Magerko
School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: nicholas.davis@gmail.com; magerko@gatech.edu

C.-P. Hsiao
College of Architecture, Georgia Institute of Technology, Atlanta, GA, USA
e-mail: chsiao9@gatech.edu

Y. Popova
Department of Cognitive Science, Case Western Reserve University, Cleveland, OH, USA
e-mail: yanna.popova@case.edu

theory of enaction claims that cognition and creativity always emerge through a real-time and improvised interaction with the environment and other agents in that environment (Varela et al. 1991; Stewart et al. 2010). While traditional theories could work to incorporate this perception-action feedback loop to model continuous improvised interaction, the enaction theory begins with the assumption that all cognition is based on this principle of improvised interactions guided by feedback from the environment. Starting from this basic assumption makes developing an enactive model of collaborative creativity and co-creation much easier due to their improvisational nature.

The overall aim of this chapter is to show how an enactive approach to computational creativity can make it easier to think about, design, and build creative computers, especially those that are able to improvise in real-time collaboration with human users. To situate and motivate our contribution, we first describe the field of computational creativity. Next, we introduce the cognitive science theory of enaction and describe creativity through its theoretical lens. Then, we present our enactive model of creativity and explain how its principles helped design "enactive" creative systems in two different domains: visual art and design.

7.2 Computational Creativity

Computational creativity is an outgrowth of artificial intelligence, cognitive science, and creativity research. It studies and builds creative systems involving different combinations of creative humans and creative computers. Making creative computers is a kind of grand challenge for the modern era of computing, and the recent efforts in computational creativity show a promising path forward. The field of computational creativity can be segmented into three broad categories that each have different motivations and goals. *Creativity support tools* augment and enhance human creativity, such as Adobe's Photoshop or Computer Aided Design tools. *Generative systems* produce creative artifacts (semi-)autonomously, such as computers that paint pictures (see Fig. 7.1) (McCorduck 1991; Colton and Wiggins 2012) or generate poetry (Colton et al. 2012). *Computer colleagues* collaborate with human users on creative tasks much like another human would (see Fig. 7.2).

Once it was established that creativity could be trained, facilitated, and measured, researchers began to develop techniques to support creativity (Smith et al. 1995; Guilford 1970; Csikszentmihalyi 1997). Initially, these techniques were procedural activities one could engage in to stimulate creativity, such as brainstorming and lateral thinking exercises (Rawlinson 1981; Bono 1970). Researchers also began developing a new class of technology referred to as creativity support tools (CSTs) (Shneiderman 2002; Shneiderman et al. 2006; Hewett et al. 2005; Carroll et al. 2009). CSTs are designed to help users explore a creative domain, record decision histories, and scaffold skills to allow and encourage users to learn expertise (Candy 1997; Shneiderman 2007).

Fig. 7.1 Art-generating computational creativity systems. *Left*: Artwork by The Painting Fool (Colton and Wiggins 2012). *Right*: Artwork by Aaron (McCorduck 1991)
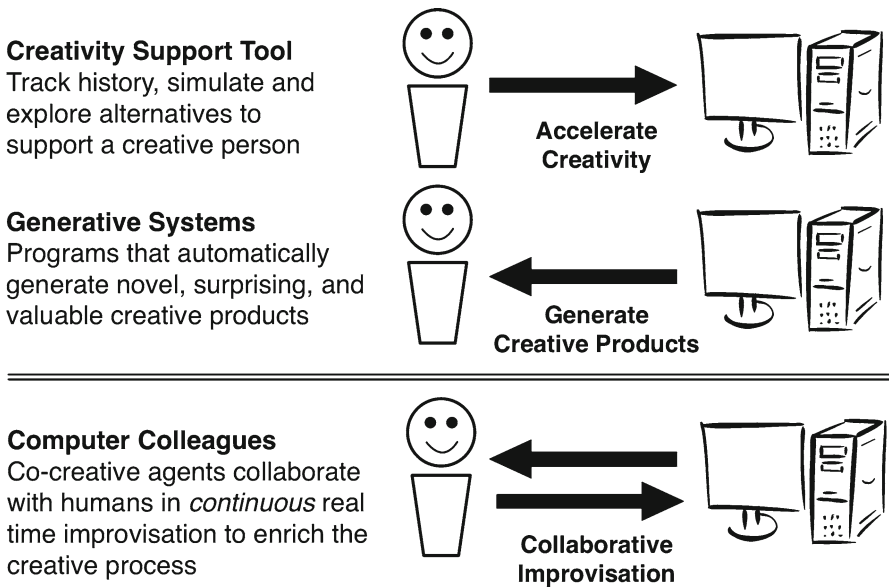

Creativity Support Tool
Track history, simulate and explore alternatives to support a creative person

Accelerate Creativity

Generative Systems
Programs that automatically generate novel, surprising, and valuable creative products

Generate Creative Products

Computer Colleagues
Co-creative agents collaborate with humans in *continuous* real time improvisation to enrich the creative process

Collaborative Improvisation

Fig. 7.2 Three approaches in the field of computational creativity

7.2.1   Creativity Support Tools

Shneiderman distinguishes creativity support tools (CSTs) from productivity support tools through three criteria: clarity of task domain and requirements, clarity of success measures, and nature of the user base (2007). Productivity support tools are designed around a clear task with known requirements, have well-defined success metrics, and are characterized by a known and relatively well-understood set of users. In contrast, CSTs often work in ill-defined domains that have unknown requirements, vague success measures, and an unpredictable user base. For example, consider productivity support tools for the well-defined goals of product supply scheduling, which include many clearly defined variables like cost metrics for shipping efficiency. Contrast this with a drawing support tool, like ShadowDraw (Lee et al. 2011) or iCanDraw (Dixon et al. 2010), that helps users learn drawing skills and inspires creativity.

Creativity support tools can take many forms. Nakakoji (2006) organizes the range of creativity support tools with three metaphors: running shoes, dumbbells, and skis (Nakakoji 2006). Running shoes improve the abilities of users to execute a creative task they are already capable of; they improve the results users get from a given set of abilities. Dumbbells support users learning about a domain to become capable without the tool itself; they build users' knowledge and abilities. Skis provide users with new experiences of creative tasks that were previously impossible; they enable new forms of execution. A contemporary text editor that highlights grammar mistakes is a running shoe; explaining why those wordings are ungrammatical makes the tool a dumbbell. Collaborative drawing tools would be a type of ski because they enable a whole new class of creative expression where the user collaborates with a computer. Nakakoji believes CSTs that introduce new creative experiences to novices will gain popularity because of the positive impact novel creative experiences can have on creative output (Nakakoji 2006).

7.2.2   Generative Computational Creativity

The class of creative systems that autonomously produce creative products is referred to here as generative computational creativity. This approach is largely inherited from AI, and it dissects human creativity into observable behaviors such as narrative, poetry, ideation, games, analogy, design, etc. These researchers then create computational models for their tightly delineated creativity module with the hope and effort to try to integrate those components with other embodied and situated aspects of creativity later.

The typical software architecture for generative computational creativity progresses as follows: The system first "reads" or interprets a large corpus of material into structured representations that it uses as its knowledge base. To make the systems more "creative," the corpus is carefully selected to lead to more interesting

combinations, such as twitter posts and news articles (Veale and Hao 2008; Colton et al. 2012). These representations form the "conceptual space" the agent traverses to find interesting combinations to produce novel output (Boden 2004). For example, a poetry-generating system might parse a news article into structured representations that can then be spliced and recombined according to hard-coded rules of poetry (meter constraints, rhyming patterns, etc.). The conceptual space itself can be restructured to reveal additional mappings and traversals within it, which is called "transformational creativity" (Boden 2004). Finally, those spaces are systematically traversed to piece together a novel creative product, which is outputted to the user. These types of creative systems typically yield bounded and discrete creative artifacts as their output. The recent 2014 International Conference on Computational Creativity, for example, was largely dominated by this approach.

Based on this distinction, a system can be referred to as generative if it does not constantly interact with its environment through both perception and action to create an artifact. Instead, it relies on building a large knowledge base from a corpus and then manipulating elements of that corpus to develop new artifacts. The "creativity" that generative systems exhibit occurs in an abstracted manipulation of symbols without a perception-action feedback loop with the environment. While the end product may resemble something we might expect of a "creative" human, we argue these systems leave out one of the most fundamental ingredients to human cognition—the environmental feedback loop.

7.2.3   Computer Colleagues

Computer colleagues are the newest and perhaps most ambitious venture in the space of computational creativity because they require a method for controlling real-time improvisational interaction with a user in addition to some mechanism for generating original creative contributions to the shared artifact. There are several options for algorithms that generate creative contributions (as discussed previously), but understanding how to get the agent to improvise in real time is difficult. A good starting point is to understand collaboration and co-creativity in humans, which is classified as multiple parties contributing to the creative process in a blended manner (Mamykina et al. 2002). It arises through collaboration where each contributor plays an equal role. Contrast this blended model with cooperation, for example, which can be modeled as a distribution of labor where the result only represents the sum of each individual contribution (Mamykina et al. 2002).

Co-creativity allows participants to improvise based on decisions of their peers. Ideas can be fused and built upon in ways that stem from the unique mix of personalities and motivations of the team members (Mamykina et al. 2002). Here, the creative product emerges through interaction and negotiation between multiple parties, and the sum is greater than individual contribution. These interaction principles can be extended to include a sufficiently creative agent that can collaborate with human users in a new kind of human-computer creativity.

Some approaches that have yielded interesting examples of computer colleagues use mimicry, structured improvisation, and shared mental models. For example, the improvisational percussion robot Shimon mimics human musicians by analyzing the rhythm and pitch of musical performances and generating synchronized melodic improvisations (Hoffman and Weinberg 2010). In practice, the human and robot develop a call-and-response interaction where each party modifies and builds on the previous contribution. Some co-creative agents use sensory input to construct mental models of agents, actions, intentions, and objects in the environment (Hodhod et al. 2012). Mental models help agents effectively structure, organize, interpret, and act on sensory data in real time, which is critical for meaningful improvisation.

Although there are only a few examples of computer colleagues today, they raise interesting questions about what it means to collaborate with a computer. These projects also point to the need for a general cognitive theory of collaboration and improvisational creativity that can be used to guide their interaction design and software architectures. We contend that enaction can fulfill this need.

7.3   The Enactive Paradigm

In the following sections, we describe how the enactive approach reframes perception into an active and dynamic process critical for participatory sensemaking, i.e., negotiating emergent actions and meaning in concert with the environment and other agents. Next, we examine the role of goals and planning in the enactive perspective. Finally, we review some sketching and design research to show evidence that enaction plays a key role in the creative process when creative individuals "think by doing."

Enactive cognition is an outgrowth of the embodiment paradigm in cognitive science. Embodiment claims cognition is largely structured by the manner in which our bodies enable us to interact with the environment (Varela et al. 1991). This approach is contrasted with earlier cognitive theories that conceptualized the mind as a machine and cognition as a complex but disembodied manipulation of symbolic representations (Newell et al. 1959). In particular, enaction emphasizes the role that perception plays in guiding and facilitating emergent action (De Jaegher 2009). A short definition of enactivism by Havelange (2010) will help summarize this distinction.

Here, cognition is no longer considered as a linear input/output sequence (as was the case in classical cognitivism) but rather in terms of a dynamic sensorimotor loop by taking into account the fact that actions themselves produce feedback effects on subsequent sensations. Action is thus no longer a simple output; it becomes actually constitutive of perception. What is perceived and recognized in perception are the invariants of the sensorimotor loops, which are inseparable from the actions of the subject.

The enactive approach takes first person experience and awareness of the cognitive agent as the starting point. It advocates for an intelligent perception and action system that pairs interesting actions and related percepts as a coupling that are stored to guide future interactions. Enaction is rooted in the notion that cognitive agents always experience reality as a continuous interaction with the world and any investigation or model should have interaction as its fundamental constituent.

7.3.1 Enactive Perception

Perception is not a passive reception of sensory data but rather an active process of visually reaching out into the environment to understand how objects can be manipulated (Gibson 1979; Noë 2004). In the enactive view, cognition is seen as a process of anticipation, assimilation, and adaptation, all of which are embedded in and contributing to a continuous process of perception and action. This type of enactive perception minimally involves a negotiation among the following factors: (1) the subject's intentional state, (2) the skills and bodily capabilities of the individual, and (3) perceptually available features of the environment that afford different actions such as size, shape, and weight (e.g., is it graspable, liftable, draggable, etc., as elaborated in Norman (1999)). Sensory data enters the cognitive system and irrelevant data is suppressed and filtered (Gaspar and McDonald 2014). Objects and details of the environment that relate to the subject's intentional goals appear to conscious perception as affordances, which can grab, direct, and guide attention and action (Norman 1999). Each time the individual physically moves through or acts upon the environment, that action changes the perceptually available features of the environment, which can reveal new relationships and opportunities for interaction.

7.3.2 Participatory Sensemaking

The enactive view accentuates the participatory nature of meaning generation, often called participatory sensemaking. Each interaction with the environment can (and often does) reveal new goals, which leads to a circuitous, rather than a linear, creative process. Creative individuals engage in a dialogue with the materials in their environment (and other agents) to define and refine creative intentions (Schön 1992).

In human daily interactions, for example, there is evidence that some form of natural coordination takes place in the shape of movement anticipation and synchronization. A good example of participatory sensemaking would be the familiar situation where you encounter someone coming from the opposite direction in a narrow passageway (De Jaegher 2009). While trying to negotiate a safe and quick passage, both participants look toward their intended path (providing a social cue) while also trying to assess the projected path of other agents. Interaction then, in the form of

coordination of movements, is the decisive factor in how quickly the individuals achieve their goal of passing each other. Rather than adopting a plan with a fixed and concrete goal to control locomotion, an enactive analysis would posit that individuals remain flexible throughout the situated action by dynamically accommodating the choice of the other agent.
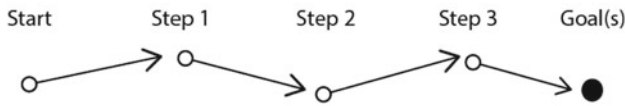
7.3.3 Goals and Directives

In the traditional view of information processing, in order to accomplish goals, an agent would follow certain steps according to a preset plan for solving the particular problems defined by concrete goals. From an enactive perspective, intelligence and creativity involve knowing how to change the flow of sensory information in order to explore possibilities for action, i.e., leaning in closer to get a better look at something. It is often simply easier to act on the environment and experiment with how different interactions affect the system than representing it in its entirety and performing symbolic processing on those representations like the information processing perspective proposes (Noë 2004). Even at the level of social interaction with an intelligent agent, an enactive approach tries to avoid postulating high-level cognitive mechanisms at the core of our intersubjective skills. The coevolution of a communicative/creative process is seen here as a gradual unfolding in real time of a dynamic system spanning a human subject, the environment, and agents within it. In this view, intentions emerge but are also transformed in and through the interaction with other agents and the environment.

Thus, instead of describing creative behavior as goal-based planning and information processing, we have adopted the enactive terminology of directives (Engel 2010). A "directive" is a loose intention that directly influences what things appear interesting or salient in the environment and how specific types of interactions might provide more information about emerging hypotheses. A directive is similar to a goal in that it can be reflected on, elaborated, and specified in more detail, but it is critically different from the current notion of "goal" in planning-based AI because it does not constitute action in any way. A directive constrains and suggests potential actions that could yield productive changes in an emergent process of sensemaking. See Fig. 7.3 for an illustration of goals compared to directives.

To illustrate the distinction between directives and goals, let us consider an example in the creative domain, such as painting a picture. Yokochi and Okada (2005) analyzed the painting process of a famous Japanese painter. He found that the artist began with a vague "directive" (our term) that is then refined and explored through interacting with the painting. Each new line adds an additional constraint and affects all the existing constraints created by previous lines. Whenever the painter decides to alter some part of the image, the enactive perspective would claim he has defined a "task" for himself. This task is similar to a goal in goal-based AI; in Fig. 7.3, tasks correspond to the small actions that serve to explore the problem space of the directive. Accomplishing a task can be modeled in an enactive manner

Conventional View of "Goals"
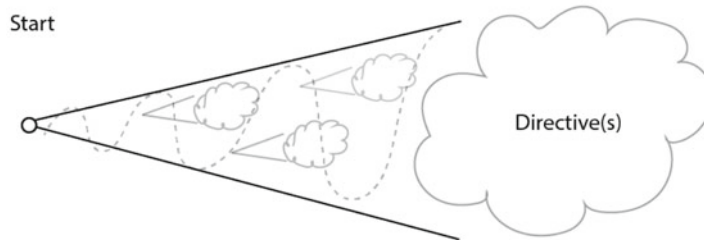

Enactive View of "Directives"

Fig. 7.3 Comparing goals and directives. Goals are linear with a series of steps whereas directives are vague and gradually refined through a process of interacting with the environment and defining tasks that explore the problem space outlined by the directive

(improvisation and affordance-based interactions) or using any number of search and planning procedures defined in goal-based AI.

Once the painter takes a step back to understand his last contribution in terms of the overall picture, however, he may find that his last contribution actually disrupted the overall balance of the piece. Although he doesn't have a specific end state for the painting in mind, one of the directives guiding his work may relate to achieving an overall balance in the composition. This directive does not tell him what contributions to make, but it helps point out inconsistencies and visual tensions that need to be addressed.

Let us suppose that the artist found five areas of the drawing that all violated his sense of balance due to his last contribution. He then selects one of those areas and defines a specific painting tasks that he predicts will help achieve balance. Once the first of those five areas is complete, the artist could take another step back and realize that his latest contribution makes the left side of the artwork look kind of like a face, which he likes. The artist might then update his overall directive to creating some kind of abstract face. Once this directive is adopted, the entire canvas is analyzed with respect to face-like features. Given this new constraint, he sees additional opportunities to change the drawing and would then select specific painting tasks that contribute toward the current directive. Here, the directive is dynamic and always evolving through interaction with the environment. The feedback offered by actually producing a change in the environment spurs new ideas and interpretations that can change the overall directive. The directive determines the constraints and affordances that are consciously available to the painter's perceptual processes.

Ultimately, it is the continuous perception-action feedback loop that actually determines actions. Instead of thinking of action as a series of behaviors executed

like scripts or plans, we can think of action as a continuous improvisation with the environment. Attention and the conscious experience of the agent become the common thread that stitches the flow of each individual action together.

Attention of the agent drives the system by changing the flow of sensory information. Depending on the current directive, the system "perceives" sensory information in different ways. At this point, the reader might ask: How can the same sensory information be perceived in different ways? If we imagine sensory input as a flow through time, we can then consider adding different "lenses" to perception to filter that sensory input in different ways. Different filters make different features of the environment salient. If they are salient enough, they will demand the attention of the individual, which might then prompt subsequent interaction. We call this filter perceptual logic because it enables a form of direct perceptual reasoning. The directive guides attention toward facets of the environment that are relevant to the current intention of the agent. The old adage "when you have a hammer everything looks like a nail" is quite illuminating to consider in this context. Once a hammer is picked up, the general directive of hammering is established, and this directive guides attention and action, which results in things being perceived in terms of their "hammerability."

To summarize the idea of a directive, a directive does not dictate action; it selects a filter for perception that (we propose) enables a perception-based reasoning process we call perceptual logic. Actions are not discrete units but rather exist as an emergent flow of interactions with the environment. Some actions are executed in service of tasks, while other actions help gain different perspectives, including changing physical location as well as changing the directive with which a scene is analyzed. This process is guided by attention and the awareness of the agent and is inherently based on the temporal flow of experience and the dynamics of interaction with the environment.

7.3.4   Enactive Creativity Thesis

To account for the emergent nature of cognition and of creativity, we can make systems that are designed from the ground up as improvisational collaborative agents. Their "intelligence" and "creativity" would then emerge organically through interacting with intelligent and creative humans. Current AI systems are good at constrained and specialized tasks, but tasks that require common sense and creativity (like collaboration and improvisation) are notoriously difficult to model computationally. Humans use what is referred to as "commonsense knowledge" to adapt their actions and understand everyday situations. The so-called commonsense problem in AI refers to the huge knowledge databases required to achieve what humans normally take for granted as common sense. Building such a large database of knowledge is notoriously difficult and labor intensive, which is one reason why a general purpose AI does not exist today. Creativity goes a few steps beyond the commonsense problem because it introduces open-ended domains that do not

necessarily have correct solutions. Collaboration further complicates this issue because it involves coordinating with other agents in a creative process. For these reasons, collaborative creativity is an extremely difficult target for traditional AI approaches.

This crack in the theoretical foundation of AI and computational creativity once seemed like a problem that would only take more computing power, larger knowledge bases, and more sophisticated machine-learning algorithms to solve. However, we think this problem reflects a larger systemic issue stemming from the basic assumptions about the nature of human cognition in AI and computational creativity. Once cognition and creativity are reframed in an enactive perspective, these hard problems become much more manageable.

Computationally creative systems employing the enactive perspective are based on a continually flowing and dynamic interaction with an environment rather than discrete actions and goal-oriented planning. An enactive investigation of creativity therefore begins at the level of perception, action, an environment, and the feedback loop that emerges during interaction. Enactive agents learn by experimentally interacting with their environment and perceiving the effects of those actions in a feedback loop, similar to a baby first learning to make sense of her senses. From this perspective, learning takes place when actions that produce a pleasing perceptual correlate (including a reaction from another agent, such as a mother cooing) are remembered as a percept-action pairing. These percept-action pairings are then repeated and built upon in an attempt to create shared meaning and experiences through participatory sensemaking, whereby agents coordinate their intentions through interaction and negotiation (Stewart et al. 2010).

The enaction theory describes creativity as a continual process whereby cognitive agents adaptively and experimentally interact with their environment through a continuous perception-action feedback loop to produce structured, organized, and meaningful interactions in an emergent process of sensemaking (or participatory sensemaking when multiple agents are collaborating). The emergent sensemaking process that results in creativity is fundamentally based on (and therefore inextricably bound to) continuous real-time interaction between an agent and its environment. During this type of emergent creativity, loose "directives" that guide actions are negotiated and fluidly defined, refined, or discarded altogether depending on how other collaborating agents and the environment respond to the agent's actions. While an enactive agent still defines directives that serve to guide actions, these directives merely constrain (rather than constitute) possible opportunities to explore in the environment.

In this process, experience, practice, and concentration help develop more nuanced and detailed percept-action couplings that afford a greater depth of interaction with the world. This means we cannot explain expertise relying exclusively on huge databases of representations manipulated in a rule-based manner (like case-based reasoning, analogical reasoning, blending, evolutionary algorithms, etc.). Experts know exactly where to look, what to look at, and when to look at it to figure out how to effectively navigate their domain of expertise. If the right information is not available, then experts know how to either restructure their sensory information

(change viewpoints) or restructure the environment (take action) in order to explore further possibilities for interaction that will in turn help evaluate emerging theories and also reveal additional actions. It is the dynamics of this feedback loop that need to be understood and modeled in order to understand the improvisation that inherently undergirds creativity.

7.3.5 *Enactive Creativity Examples*

The literature on creativity provides evidence supporting the enactive perspective with research on "thinking by doing." There is a multitude of evidence demonstrating how both representational and nonrepresentational artists plan their artworks using sketches, studies, and other ways to simulate artistic alternatives (Mace and Ward 2002). Sketching reduces cognitive load and facilitates perceptually based reasoning (Schön 1992). In many creative domains, individuals generate vague ideas and then use some form of sketch or prototyping activities to creatively explore, evaluate, and refine artistic intentions (Davis et al. 2011). Sketching allows creative individuals to think by doing. When an action or idea is materialized in some way, the perceptual system is rewarded with richer data than pure mental simulations and abstract reasoning. Additionally, cognitive resources that would have been used to simulate the action (i.e., consciously visualizing the situation) are now freed for other tasks such as interpretation and analysis (Shneiderman 2007).

7.3.5.1 Architectural Design

One obvious example of using sketch to "think by doing" can be found in the task of planning the spatial configurations in the architectural design process. As addressed above, generating an entire artifact with all of its details directly from the mind is virtually impossible for a designer (Schön 1992). Instead, designers experience these improvised real-time adjustments in the design procedures with the tools and materials they are using. When starting the design process, designers choose different materials, tools, and media to present the initial ideas from their minds to explore the constraints of their problem (Schön 1992). When they interact with these tools, they might need to adjust their actions in order to achieve their needs. For instance, when drawing a sketch to study the forms, they may need to constantly adjust the "next steps" in order to solve the design constraints, such as not enough space, too long, too much curvature, etc.

Figure 7.4 illustrates a typical spatial plan of a student center in a bubble diagram. Since the plan entails many spaces, the designers would have to write down all the space names so that related spaces are located next to each other. They would also use arrows to represent the main circulation paths between two spaces. Each time a new space is added or an arrow is inserted, the designer's flow of sensory information changes and they might discover new problems or opportunities that were not apparent before (Suwa and Tversky 1997). Sketching facilitates their
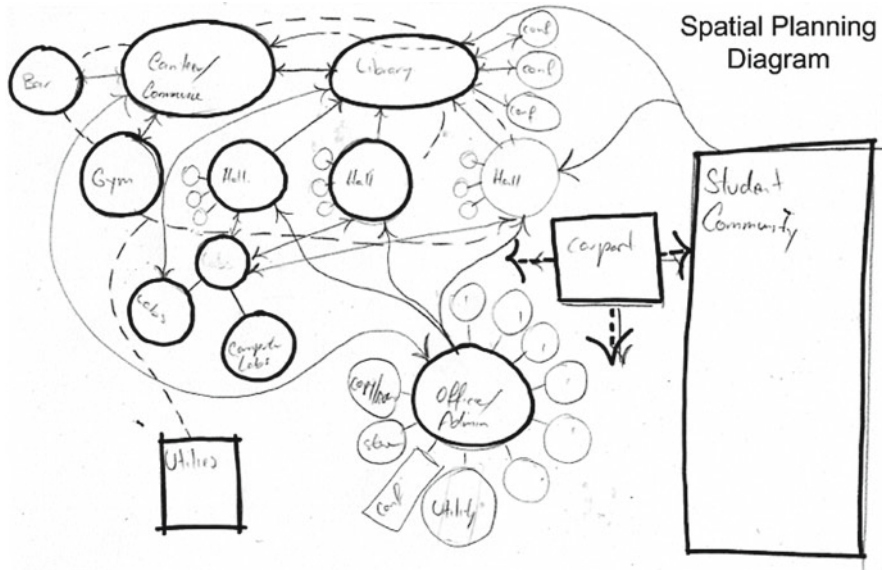
Fig. 7.4 Spatial layout of a school student center design (Courtesy of Kyle Doggett)

creativity and reasoning process through a dynamic perception-action feedback loop whereby new meanings are gradually constructed through a negotiation with the design materials (i.e., sketch, physical models, computational models, etc.).

Experienced designers also change the granularity of their perception to reason about sketches at different levels. When focusing on individual details, an architect might imagine how a particular corridor might feel to walk through. Then, they could shift to a global perspective that considers the overall theme and consistency of the whole building design (Suwa and Tversky 1997; Goldschmidt 1991).

7.3.5.2 Musical Performance

The enactive nature of creativity can also be seen in live musical performance. A classical musician, for example, a trumpet player, will need to feel the acoustic effects in a concert hall before his performances. For instance, he may extend the ending of a sound in a concert hall that has a "dry" acoustic effect. We propose the expert trumpet player has a well-established set of percept-action pairings (creating his expert perceptual logic) that have to be tuned to the particular performance space because the actions he will take in the performance will result in a slightly different perceptual feedback process than his normal practice space. Thus, he has to actively feel and explore the sounds of the space to align his perceptual logic with the specifics of the exact situation. Furthermore, during performance, he will also listen to the mixture of his trumpet sound with other sounds to make real-time adjustments to achieve the desired general effective (i.e., the directive, such as playing a "sad" tune).

7.3.5.3 Visual Art

The enactive nature of creativity in visual art is demonstrated well by the findings showing that expert artists often step away from their paintings to gain a new perspective (Yokochi and Okada 2005). Here, enaction would claim that expert artists have acquired percept-action pairings that constitute experiential knowledge: Altering the flow of sensory information can reveal additional possibilities for action. The percept-action coupling is moving the body (actions) and gaining different viewpoints (percepts). There is no preset specific goal driving the artist's decision to step back, and there is not a "step-back-and-think script" the artist executes at predefined times. Instead, there might be some open questions about how to interact with different regions of the artwork and a vague intention to address those concerns. Stepping back helps think about how interacting with those areas might affect the overall vague intention. The "creative" behavior of stepping back is actually an emergent by-product of how cognition and creativity work. The fact that the artist stepped back (her behavior) is therefore not as important as why she stepped back, i.e., how she knew that stepping back was the right thing to do. An expert is an expert precisely because she knows how to direct her attention and manipulate the flow of sensory information through interactions with the environment to explore and evaluate possibilities for further action.

The domain-independent examples above provide evidence that creativity does not only come from executing planned steps and actions but emerges through improvisational micro interactions between the human and the surrounding environments, including other humans, tools, and, most importantly, the continuous results generated during the percept-action feedback loop. We consider these interaction processes as an improvised interaction processes. Humans often experience the results from unplanned micro interactions that match or mismatch their expectations, which will then become perceptual logic for future interactions. We argue that this enactive feature of cognition is fundamental to understanding how to understand human creativity and also build computer colleagues.

7.3.6 Enactive Model of Creativity

The argument here is that the traditional cognitive science theories used by AI are inadequate to explain the entirety of human creativity (and cognition more broadly) and should thus be supplemented, augmented, or potentially replaced entirely with an enactive conceptualization of cognition. In the enactive view, cognition (including creativity) is inherently composed of a continuous interaction with the environment and other agents in that environment to adapt and thrive (Stewart et al. 2010); it is improvisational and ever changing based on the demands and opportunities of the moment. The enactive view encapsulates the embodied, situated, distributed cognition perspectives that have recently gained popularity (Suchman 1986; Hutchins 1995). From this view, cognition is not inherently goal-based planning procedures, as the search and planning-centric approaches in AI suggest. Although

we certainly construct plans to try to organize our interactions with the environment, they are never constitutive of the actual creative process, which is enacted in concert with feedback from the environment. We cannot cut off this real-time interaction feedback loop with the environment in any way if we hope to create a realistic model of creativity and cognition.

7.3.6.1   Model Description

We first explain the visual conventions of the enactive model of creativity and describe how it can be applied to model creative cognition through time. Then, we describe in detail a new concept derived from our model called perceptual logic, which is a perceptual filter that highlights relevant affordances in the environment while suppressing irrelevant affordances. Next, we explain how modulating perceptual logic leads to different ways of seeing and interacting with the world in a way that can account for the diverse array of human creative behavior.

In the enactive model of creativity (see Fig. 7.5), the awareness of the agent is represented by the vertical rectangle situated on a spectrum of cognition, which essentially means that the agent is "aware" of what is perceived and its current intention. Perception is constituted partly by the mental model the agent has constructed for the current situation (top-down cognition) as well as the sensory input coming from the environment (bottom-up cognition) (Gibson 1979; Glenberg 1997; Varela et al. 1991; Stewart et al. 2010; Gabora 2010).
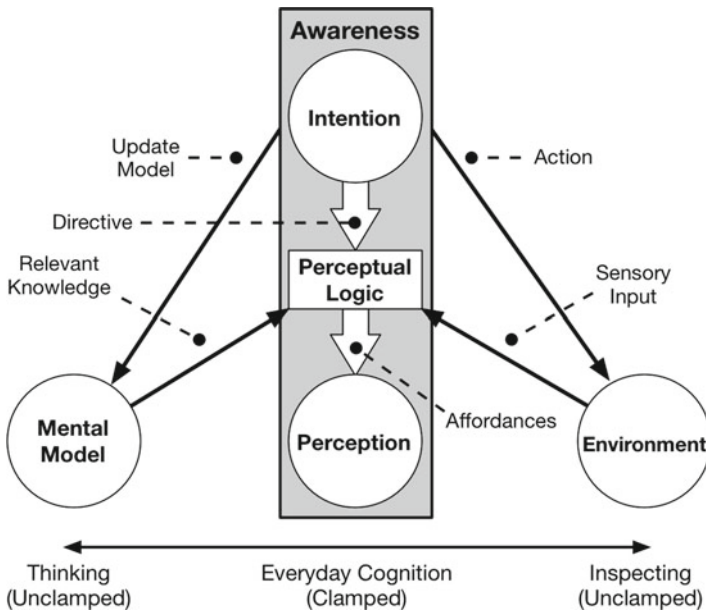

Fig. 7.5 Enactive model of creativity

To get a sense of the intended dynamism of this model, imagine the entire "awareness" rectangle as one unit that can shift to the left or right on the cognitive continuum as a function of the agent's concentration. Routine actions only require minimal thought and a limited amount of highly relevant sensory data. The enactive model of routine actions, such as driving, would be visually depicted by having the awareness rectangle resting at equilibrium in the center of the spectrum with small deviations to the left to update and revise strategy and deviations to the right to interactively evaluate those ideas.

To simulate bounds on working memory, the agent only has a limited amount of cognitive resources. These resources are used through a process of directed attention, i.e., concentration. During this simulated form of concentration, agents devote their attention to reflecting on the situation (building more detailed mental models, running complex mental simulations, etc.) or acting in a deliberate and interactive manner to inspect the world.

If the agent is performing an unfamiliar task, however, cognitive resources are recruited to actively build a mental model of the situation, which requires performing experimental interactions, closely examining the results in the environment, and then updating the mental model in a slower global model of perceptual logic. Initially, novices have to think a lot about what they are doing, which means they are using a lot of the previous attention resources to build up a cognitive model by performing micro experiments interacting with the world to hypothesize about this particular domain. As novices build up this model, they begin to interact without having to pay as much attention to what they are doing. The enactive model claims this happens because the experienced individual is able to use the new perceptual logic to filter irrelevant sensory details and operate effectively with minimum conscious supervision of a task (see Fig. 7.6 for an illustration of different layers of perceptual logic).

7.3.6.2 Perceptual Logic

According to the enactive model of creativity, the contents of perception vary based on an individual's position on this continuum of cognition (Glenberg 1997). As individuals deviate from the equilibrium in the center of the spectrum, perception becomes partially "unclamped" (a term coming from Glenberg's (1997) theory of memory) which loosens semantic constraints on sensory input and memory. Different points on the cognitive spectrum result in a unique perceptual logic that is used to intelligently perceive affordances in the environment. The enactive approach in cognitive science describes the "intelligence" of perception in a theoretical sense, but operationalizing the theory required explaining the implicit black box mechanism that makes perception "intelligent." The mechanism basically serves to filter all possible affordances and present only relevant affordances to conscious perception.

The enactive approach proposes that perceptual intelligence arises through the formation of percept-action pairings that are chunked and internalized for quick

retrieval (Noë 2004). Perceptual logic is a proposed cognitive mechanism that filters sensory data, identifies relevant percept-action pairings, and presents these percept-action pairings as affordances to perception. Perceptual logic performs a similar role as the "simulator" in Perceptual Symbol Systems (Barsalou 1999). The simulator activates all the associated neural correlates related to a percept, including the various ways it can be interacted with based on experiential knowledge and physical characteristics.

7.3.6.3   Clamping Perception

Research indicates that perception filters irrelevant sensory input to reduce distractions and facilitate everyday cognition (Gaspar and McDonald 2014). When the agent is engaged in a routine task and following well-established affordances, sensory data is "clamped" to filter out unnecessary details and unconventional ways of seeing objects (Glenberg 1997). Everyday cognition is represented in EMC by situating the awareness rectangle in the center of the spectrum of cognition, creating a point of equilibrium. Shifting to either the left or right on this spectrum requires the agent to either concentrate on the details of her mental model or closely inspect details in the environment. At equilibrium, perception is clamped to a combination of sensory input and cognitive input that optimizes routine interactions (Glenberg 1997). When minor problems arise, such as small improvisational adjustments to the action based on environmental feedback, this equilibrium is slightly perturbed. The agent could generate various alternative actions by thinking (moving slightly left on the spectrum) and explore various ideas by interacting with the environment (moving slightly right on the spectrum).

7.3.6.4   Unclamping Perception

If there is a severe disruption to the current task (e.g., a great new idea, distraction, or some kind of failure), it might become necessary to disengage from the current task to reevaluate the situation. When an individual "disengages" from a task, perception becomes "unclamped" and attention shifts to thinking and simulating solutions (moving far left on spectrum) and closely examining the detail of the environment to discover new affordances (moving far right on the spectrum). The degree of concentration devoted to thinking about or acting on the environment determines how far, in either direction, awareness is situated on the spectrum of cognition. At the extreme left of the continuum (thinking) would be closing one's eyes to try to think deeply about a topic, which removes most sensory input from perception altogether. At the extreme right of the continuum (inspecting) would be an individual fully concentrated on acting skillfully, carefully, and deliberately on the environment.

7.3.6.5 Modulating Semantic Constraints

During these periods of disengaged evaluation, EMC proposes that the semantic constraints for recalling associated ideas from memory and interpreting elements in the environment become "unclamped" to enable reconceptualization. Unclamping semantic constraints helps overcome functional fixedness, which is a phenomenon where individuals have trouble dissociating objects from their entrenched meaning during insight problem-solving (Adamson 1952).

In the cognitive science literature, the abovementioned type of meaning reassignment is referred to as a *conceptual shift* (Nersessian 2008). Colloquially termed the Eureka! or Aha! moment, conceptual shifts occur when two separate knowledge domains are connected in the mind (Boden 2004; Nersessian 2008). It is often partially or wholly responsible for insights that lead to creative discoveries and solutions. The enactive model suggests that conceptual shifts and creative reconceptualizations are made possible by unclamping perception, thereby allowing new meanings to be associated with objects and concepts.

Interestingly, this model identifies an important role for distraction in the creative process. Distraction is one way to prompt an individual to disengage from everyday cognition. In abstract art, for example, unfinished segments of the artwork (or unexpected contributions from a collaborator) may distract the artist while they are drawing. These newly discovered areas might not align with the artist's current intention. As a result, the artist might want to resolve that tension by drawing additional lines, which can catalyze the creative process. However, too many distractions might frustrate the artist.

Now that we have introduced enaction and presented the enactive model of creativity, we will describe how this model was helpful in designing two computer colleagues in the domains of visual art and design.

7.4 Building Co-creative Agents with the Enactive Model

The enactive model of creativity served as a productive framework to design co-creative agents because it enables agents to interactively adapt their perceptual reasoning strategies and creative behavior to that of the user, which increases the probability the user will find the contributions of the system meaningful and creatively engaging.

7.4.1 Layers of Perceptual Logic

There are three layers of perceptual logic in the enactive model of creativity (local, regional, and global) that are determined by the position of awareness on the spectrum of cognition (see Fig. 7.6). Each successive layer of perceptual logic considers a larger portion of the creative artifact (i.e., more sensory data) at a higher level of
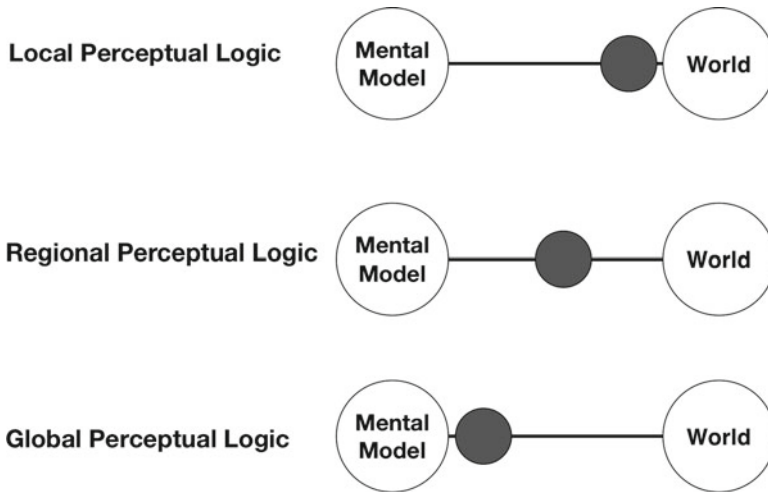
Fig. 7.6 The layers of perceptual logic: The position of awareness (*gray ball*) on the spectrum of cognition corresponds to the layer of perceptual logic the system uses

conceptual abstraction (global being the most complex). Since each layer is more complex than the next, we found the most effective implementation strategy to be implementing them progressively in stages starting with the most basic local layer of perceptual logic.

Local perceptual logic considers granular details of the user's contributions, such as individual lines added to a drawing. Regional perceptual logic, on the other hand, groups the user's inputs into regions and containers based on principles of gestalt grouping, such as proximity, similarity, common fate, and continuity (Arnheim 1954). The principles of gestalt grouping were encoded into this layer of perceptual logic to provide a means for the system to begin to make sense of creative contributions in a similar way as humans.

Global perceptual logic considers the creative artifact as a whole, like when an artist takes a step back from their painting. This form of perceptual logic considers the relationship between the different regions of the drawing to analyze the overall composition. When this perceptual logic is applied, the system may decide to completely decouple its contribution from the human's recent input, i.e., it can select non-active regions of the artifact on which to operate if those regions present significant creative opportunities. For example, a drawing system might examine the overall composition and determine that the left side of the drawing is imbalanced because it has significantly less lines overall than the right side of the drawing. The system employs global perceptual logic to reason about the drawing as a whole and set a directive of "do work on the left-hand side of the drawing." After this directive is determined, the system would then employ either regional or local perceptual logic to determine the exact lines to draw on the left-hand side of the page. The directive therefore constrains the possible actions the system could potentially take and guides interaction going forward, but it does not determine actions in any way, which is the critical difference between directives and goals.

7.4.2 Drawing Apprentice

The Drawing Apprentice is a co-creative agent that collaborates with human users to draw abstract artworks on a digital canvas in real time (Davis et al. 2014). The system improvises with users in a turn-taking manner. First, the user draws a line. The system then reacts with a line of its own. The system analyzes the user's lines and drawing behavior (i.e., line length, speed, time between strokes, location, etc.) through time to construct a directive. This directive guides how the agent perceives its environment (lines) by applying one of the three layers of perceptual logic that each consider different scales of the drawing (i.e., local, regional, and global). Local perceptual logic modifies individual lines (i.e., mirror, translate, scale, trace, shade, etc.) and redraws them. Regional perceptual logic employs gestalt principles to group lines into regions that can be modified in a similar way as individual lines. Finally, the agent can consider the relationship between groups to evaluate the over-all composition, such as balance. The agent doesn't have any pre-encoded drawing algorithms, per se. It only has the ability to direct its attention, perceive the user's lines, and manipulate and interact with those lines according to its perceptual logic. The program will be provided with some perceptual rules of gestalt grouping to inform perception how to group sensory input into larger gestalt wholes (i.e., principles of perceptual grouping: good continuity, closure, proximity, flow, etc.) that allow the system to build its own knowledge base through its experience collaborating with artists (Fig. 7.7).

7.4.3 Multiple Sets of Perceptual Logic

The argument we have built in this chapter contends that experts gradually develop perceptual logic that enables them to intelligently perceive their environment to navigate specific situations. When a creative expert attempts to accomplish their creation process on a creativity support tool, like a designer using a traditional CAD tool, they have to acquire a completely different set of perceptual logic relating to how to navigate the interface and accomplish tasks. Users have to alternate between these sets of perceptual logic when they interact with creativity support tools, which can take users out of the immersive and interactive flow that the enactive model of creativity proposes is critical for facilitating creativity. As a result, people often use CAD tools at late stages in the design process to finalize their design, instead of using them to facilitate creative thinking and exploration early in the design process. One overarching design principle of an enactive approach is to design interactions as conversations, where each party tries to understand and build meaning through negotiation and feedback over time. In a conversation, each person actively works to understand what was said and respond appropriately. This suggests that creativity support tools might develop a dynamic model of the user over time based on their interactions and behaviors such that we might understand what type of perceptual
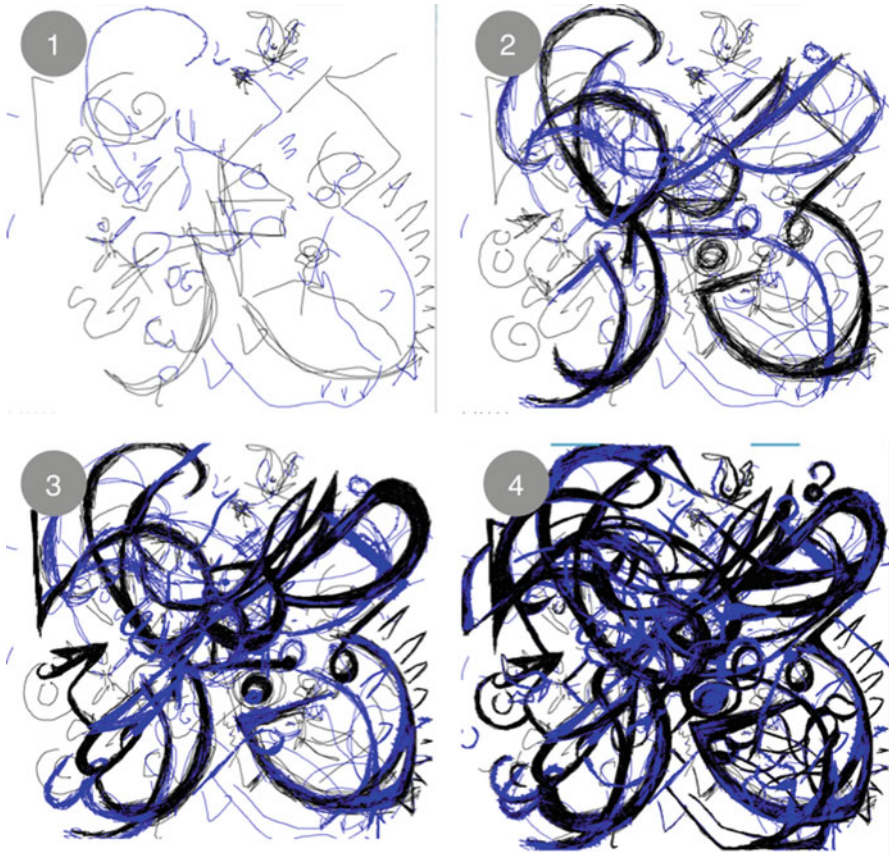
Fig. 7.7 Drawing Apprentice collaboration. User's lines are *black*; AI agent's lines are *blue* (Color figure online)

logic and creative strategy the user is currently employing and offer the right tools at the right time.

7.4.4   Solid Sketch

Solid Sketch is an example of a CST that utilizes the concepts we describe in the previous section. It is a sketch program for 3D model creation that constantly observes the user's sketch inputs and reacts in real time based on the context determined by the previous and surrounding sketches. The enactive model of creativity serves two roles in this prototype. One is to help the system understand the perceptual logic the user employs throughout their creative process. The other use is to facilitate natural interactions when designing the prototype. For the first purpose,
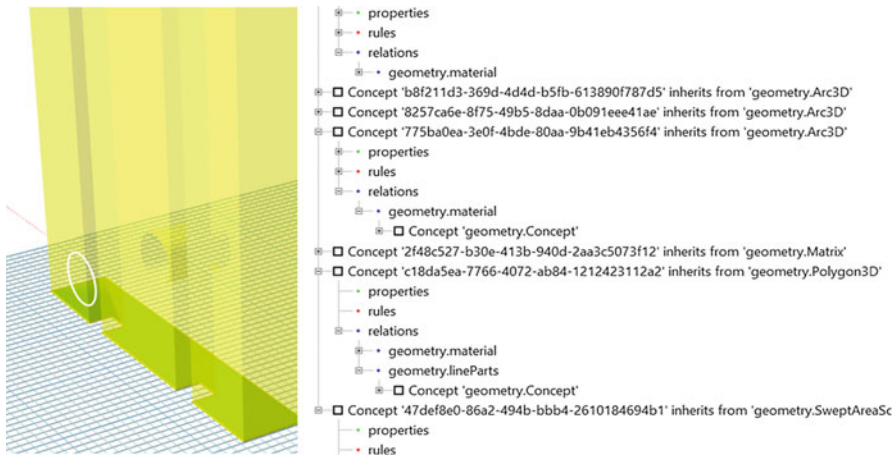
Fig. 7.8 *Left*: A simple 3D model done with Solid Sketch. *Right*: The system interprets the human natural sketch into parametric information

the program uses the enactive model of creativity to construct cognitive models of how humans construct the entire 3D model from sketches at different levels, e.g., local, regional, and global. For the system, its local perceptual logic tries to understand the relationship between the geometry, such as the angle between two sketch lines. Regional perceptual logic attempts to compose nearby sketch lines into coherent part of the model. Global perceptual logic composes those regional perceptual logic groupings into a meaningful overall model. The second use of the enactive model is to facilitate the conversation like creation process instead of having users to execute commands explicitly one by one, such as the traditional CSTs that require users to execute commands and input complicated equations explicitly. The enactive agent in Solid Sketch sits in the background, perceives the user's actions, interprets his intentions, and leverages its understanding of the user's intention to help the user achieve their current goal. The final products after interacting with the system will include not only a 3D model but also a set of parametric rules that describe how the user created the model (Fig. 7.8).

7.5   Conclusions

Computational creativity has the potential to radically change what it means to interact with computers. However, in order to reach its full potential, the field needs a cognitive theory of creativity that accounts for the enactive nature of creativity, including improvisation, collaboration, and a tight feedback loop with the environment. In this chapter, we provided a brief summary of the current state of computational creativity and pointed out the shortcomings of the traditional information processing view of cognition. We argued that the new cognitive science paradigm of

enaction provides a helpful way to reframe creativity and potentially solve some of the long-standing hard problems that both artificial intelligence and computational creativity face. The theory of enaction was used to describe creativity in design, music, and visual art to show its potential for generalizability and descriptive power. We also presented the enactive model of creativity that formalized the enaction theory in a computational model. Finally, we describe how the enactive model of creativity was helpful in designing two computer colleagues, one in the domain of visual art and the other in the domain of design. The primary design principle of the enactive model of creativity is to design interactions like a conversation where each party tries to make sense of contributions and respond appropriately given the history of interaction.

References

Adamson RE (1952) Functional fixedness as related to problem solving: a repetition of three experiments. J Exp Psychol 44(4):288–291

Arnheim R (1954) Art and visual perception. University of California Press, Oakland

Barsalou LW (1999) Perceptual symbol systems. Behav Brain Sci 22(04):637–660

Biles JA (2003) GenJam in perspective: a tentative taxonomy for GA music and art systems. Leonardo 36(1):43–45

Boden MA (2004) The creative mind: myths and mechanisms. Psychology Press, New York

Bono ED (1970) Lateral thinking: a textbook of creativity. Ward Lock Educational/El pensamiento lateral: manual de creatividad, Londres/Versión española

Candy L (1997) Computers and creativity support: knowledge, visualisation and collaboration. Knowl-Based Syst 10(1):3–13

Carroll EA, Latulipe C, Fung R, Terry M (2009) Creativity factor evaluation: towards a standardized survey metric for creativity support. In: Proceedings of the seventh ACM conference on creativity and cognition. ACM, New York, pp 127–136

Colton S, Wiggins GA (2012) Computational creativity: the final frontier? In: Proceedings of European conference on artificial intelligence, pp 21–26

Colton S, Goodwin J, Veale T (2012) Full face poetry generation. In: Proceedings of the third international conference on computational creativity. International Conference on Computational Creativity, pp 95–102

Csikszentmihalyi M (1997) Flow and the psychology of discovery and invention. HarperPerennial, New York

Davis N, Li B, O'Neill B, Riedl M, Nitsche M (2011) Distributed creative cognition in digital filmmaking. In: Proceedings of the 8th ACM conference on creativity and cognition. ACM, New York, pp 207–216

Davis N, Popova Y, Sysoev I, Hsiao CP, Zhang D, Magerko B (2014) Building artistic computer colleagues with an enactive model of creativity. In: Proceedings of the fifth international conference on computational creativity. The International Association for Computational Creativity, pp 38–45

De Jaegher H (2009) Social understanding through direct perception? Yes, by interacting. Conscious Cogn 18(2):535–542

Dixon D, Prasad M, Hammond T (2010) iCanDraw: using sketch recognition and corrective feedback to assist a user in drawing human faces. In: Proceedings of the SIGCHI conference on human factors in computing systems. ACM, New York, pp 897–906

Engel AK (2010) Directive minds: how dynamics shapes cognition. In: Enaction: towards a new paradigm for cognitive science. MIT Press, Cambridge, MA, pp 219–243

Gabora L (2010) Revenge of the "neurds": characterizing creative thought in terms of the structure and dynamics of memory. Creat Res J 22(1):1–13

Gaspar JM, McDonald JJ (2014) Suppression of salient objects prevents distraction in visual search. J Neurosci 34(16):5658–5666

Gibson JJ (1979) The ecological approach to visual perception. Lawrence Erlbaum Associates, Hillsdale

Glenberg AM (1997) What memory is for: creating meaning in the service of action. Behav Brain Sci 20(01):41–50

Goldschmidt G (1991) The dialectics of sketching. Creat Res J 4(2):123–143

Guilford JP (1970) Creativity: retrospect and prospect*. J Creat Behav 4(3):149–168

Havelange V (2010) The ontological constitution of cognition and the epistemological constitution of cognitive science: phenomenology, enaction and technology. In: Enaction: towards a new paradigm for cognitive science. MIT Press, Cambridge, MA, pp 335–360

Hewett T, Czerwinski M, Terry M, Nunamaker J, Candy L, Kules B, Sylvan E (2005) Creativity support tool evaluation methods and metrics. Creativity Support Tools, 10–24

Hodhod R, Piplica A, Magerko B (2012) A formal architecture of shared mental models for computational improvisational agents. In: Intelligent virtual agents. Springer, Berlin/Heidelberg, pp 440–446

Hoffman G, Weinberg G (2010) Shimon: an interactive improvisational robotic marimba player. In: CHI'10 extended abstracts on human factors in computing systems. ACM, New York, pp 3097–3102

Hutchins E (1995) Cognition in the wild. The MIT Press, Cambridge, MA

Lee YJ, Zitnick CL, Cohen MF (2011) ShadowDraw: real-time user guidance for freehand drawing. ACM Trans Graph 30(4):27:1–27:10. doi:10.1145/2010324.1964922

Lubart T (2005) How can computers be partners in the creative process: classification and commentary on the special issue. Int J Hum Comput Stud 63(4):365–369

Mace MA, Ward T (2002) Modeling the creative process: a grounded theory analysis of creativity in the domain of art making. Creat Res J 14(2):179–192

Mamykina L, Candy L, Edmonds E (2002) Collaborative creativity. Commun ACM 45(10):96–99

McCorduck P (1991) Aaron's code. WH Freeman & Co., New York

Nakakoji K (2006) Meanings of tools, support, and uses for creative design processes. In: International design research symposium '06, Seoul

Nersessian N (2008) Creating scientific concepts. MIT Press, Cambridge, MA

Newell A, Shaw JC, Simon HA (1959) The processes of creative thinking. Rand Corporation, Santa Monica

Noë A (2004) Action in perception. MIT Press, Cambridge, MA

Norman DA (1999) Affordance, conventions, and design. Interactions 6(3):38–43

Rawlinson JG (1981) Creative thinking and brainstorming. Gower, Farnborough

Schön DA (1992) Designing as reflective conversation with the materials of a design situation. Knowl-Based Syst 5(1):3–14

Shneiderman B (2002) Creativity support tools. Commun ACM 45(10):116–120

Shneiderman B (2007) Creativity support tools: accelerating discovery and innovation. Commun ACM 50(12):20–32. doi:10.1145/1323688.1323689

Shneiderman B, Fischer G, Czerwinski M, Resnick M, Myers B, Candy L, Terry M (2006) Creativity support tools: report from a US National Science Foundation sponsored workshop. Int J Hum Comput Interact 20(2):61–77

Smith SM, Ward TB, Finke RA (eds) (1995) The creative cognition approach. MIT Press, Cambridge, MA

Stewart JR, Gapenne O, Di Paolo EA (eds) (2010) Enaction: toward a new paradigm for cognitive science. MIT Press, Cambridge, MA

Suchman L (1986) Plans and situated actions. Cambridge University, New York

Suwa M, Tversky B (1997) What do architects and students perceive in their design sketches? A protocol analysis. Des Stud 18(4):385–403

Varela FJ, Rosch E, Thompson E (1991) The embodied mind: cognitive science and human experi-
    ence. MIT Press, Cambridge, MA
Veale T, Hao Y (2008) A fluid knowledge representation for understanding and generating creative
    metaphors. In: Proceedings of the 22nd international conference on computational linguistics,
    vol 1. Association for Computational Linguistics, Stroudsburg, PA, pp 945–952
Yokochi S, Okada T (2005) Creative cognitive process of art making: a field study of a traditional
    Chinese ink painter. Creat Res J 17(2–3):241–255
Zook A, Magerko B, Riedl M (2011) Formally modeling pretend object play. In: Proceedings of
    the 8th ACM conference on creativity and cognition. ACM, New York, pp 147–156