Intrinsically Motivated Creative Search

One can argue that any process capable of creating something truly new has to be selectional: it has to actively try alternatives, process resulting evaluations, and use some selection mechanism to guide behavior toward the better alternatives. Implicit in this view is a creative search paradigm that can be considered to be a sophisticated form of trial-and-error learning. Such creative search often takes place in very high-dimensional spaces in which evaluative information may only be received upon completion of a creative endeavor, or episode. Successful creative search must therefore exploit the structure of these spaces to constrain the search space, and use self-generated, intermediate feedback as a surrogate for the sparse reward signals provided by the environment. We argue that a successful creative enterprise emerges from the interaction of a sufficiently smart generator and a sufficiently smart evaluator, which perform these functions respectively. The generator must actively discover structure in the environment and use knowledge of this structure to propose what it believes to be highly valued trajectories through the creative search space. Complementarily, the evaluator must use its knowledge of past creative endeavors and their accompanying payoffs to accept or reject the plans proposed by the generator and to guide the intermediate steps of the enterprise by assigning credit appropriately along the way. We propose a computational system that implements this interaction and intend to show creative performance in a simulated domain.

By virtue of its strong theoretical foundations and its ties to existing research on credit assignment and sequential decision making, reinforcement learning is a natural choice for a computational framework in which our approach to creative search can be formalized as an optimization problem and analyzed rigorously. We choose the factored Markov Decision Process (MDP) as our environment model, since it affords greater exploitation of structure than standard MDPs. We also make use of existing research into intrinsically motivated reinforcement learning to guide our smart generator to actively discover the structure of its environment and propose novel creative endeavors. This is achieved through definition of an intrinsic reward function that rewards both discovery of structure directly as well as discovery of novel and incongruent situations in the environment. The success and originality of our approach is contingent upon the creative agent discovering behavioral building blocks that allow it to change the values of variables in its environment reliably by learning temporally and spatially abstract skills. Using existing methods for hierarchical decomposition of factored MDPs given known structure, our system will actively discover structure in its environment, generate and learn abstract skills for exploiting this structure, and set goals that use these skills to reach novel or incongruent situations in its environment. The hierarchical nature of the skill-learning system will allow the agent to set incrementally more difficult goals for itself and thus potentially propose more creative experiments to execute as it continues to explore.

Our test environment is an artificial chemistry laboratory in which an agent can combine compounds of varying chemical structure and physical properties in different ways to produce new compounds with differing structures and properties. The design of the domain allows for creative search in that manipulation of one physical property should provoke manipulation of the same property in different contexts (e.g., in the context of different combinations of other properties), which may not have been previously observed.