As Eric Baum says in “What is Thought?”, reinforcement learning is essential to intelligence. A brain senses and acts in the world, and learns behaviors reinforced by values for distinguishing good and bad outcomes. The brain learns a simulation model for tracing cause and effect relations between behaviors and outcomes – that is, for solving the credit assignment problem. Reason and high level representations of sense information are part of this simulation model, and language is a representation of the model for exchange with other brains. Thus the simulation model and its role in reinforcement learning provide a context for integrating different AI subfields.

Brains can learn simulation models as internal behaviors that predict sense information, reinforced by predictive accuracy. For example, learning to predict short-term changes to visual information based on body motion may be the basis for learning 3-D allocentric representations of vision. Learning to predict longer-term changes to visual information, sometimes in response to the brain's motor behaviors, may be the basis for learning to partition the visual field into objects, learning to classify those objects, and learning to model behaviors of object classes.

A brain design may be partitioned into interacting learning processes, each defined by a set of inputs (some sensory, some from other brain processes), an internal representation, a set of outputs (some motor, some to other brain processes), and a reinforcement value. The designs of these brain processes would be driven by knowledge and techniques from different AI subfields, cast as learning processes (this may be difficult for some research, but my assumption is that learning is fundamental to all intelligent behavior). For example, one process may take raw vision and body motor controls as inputs, produce a 3-D allocentric representation as output, and reinforce based on the accuracy of predicting raw visual input using the 3-D allocentric representation and body motion. Vision research can constrain the mapping from raw vision to 3-D allocentric representation in order to increase learning efficiency, and probably also dictate the need for inputs from other vision processes (neuroscience suggests complex connections among vision processes). Processes reinforced by short-term predictive accuracy may produce useful input to processes making longer-term predictions. Predictive processes may help trace cause and effect relations between behaviors and rewards in other processes.

Integrating language research will be difficult because language represents most of the simulation model, so language processes must connect to details in most other brain processes. This complexity has been previously expressed by those arguing for the need to solve the symbol grounding problem in order to create intelligent language behavior. One particularly difficult aspect of language is its role as a shortcut for brains to learn parts of their simulation models from other brains. As with other social processes, most reinforcement for learning language behavior must come from these other “teacher” brains.

The brain’s simulation model is used for planning. In “Reinforcement Learning”, Sutton and Barto explain how planning can help solve the credit assignment problem. At a high cognitive level, consider how chess players and mathematicians learn from the successes and failures of their plans. Their successful consciously planned behaviors are learned as fast, unconscious responses, available as behavior elements in future plans. A brain design needs processes that implement such high-level planning and learning.