Neural circuit policies

Prof. Dr. Radu Grosu, TU Wien

A central goal of artificial intelligence is to design algorithms that are both generalisable and interpretable. We combine brain-inspired neural computation principles and scalable deep learning architectures to design compact neural controllers for task-specific compartments of a full-stack autonomous vehicle control system. We show that a single algorithm with 19 control neurons, connecting 32 encapsulated input features to outputs by 253 synapses, learns to map high-dimensional inputs into steering commands. This system shows superior generalisability, interpretability and robustness compared with orders-of-magnitude larger black-box learning systems. The obtained neural agents enable high-fidelity autonomy for task-specific parts of a complex autonomous system.