Challenges

- Planning
  - MCTS
  - Learning a model
  - C & M architectures
    - Making the world differentiable (1990)
    - Online algorithm for dynamic RL and Planning in Reactive environments (1990)
    - RL in Markovian and non-Markovian environments (1991)
    - On Learning To Think (2015)
    - World Models (2018)
  - One Big Net (2018)
- Temporal abstractions
  - Hierarchical RL
    - Learning to generate artificial fovea trajectories for target detection (1991)
    - Learning to generate sub-goals for action sequences (1991)
    - Planning simple trajectories using neural subgoal generators (1992)
    - HQ-Learning (1997)
    - Options framework
    - STRAW
    - Feudal Networks
    - HIRO
- Exploration
  - Randomness / Noise
    - Model
    - Policy
  - Uncertainty estimates
  - Seeking novelty / Intrinsic Motivation
    - A possibility for implementing curiosity and boredom in model-building neural controllers (1991)
    - Curious model-building control systems (1991)
    - Reinforcement-driven information acquisition in non-deterministic environments (1995)
    - What's interesting? (1997)
    - Exploring the Predictable (2002)
    - Continual curiosity-driven skill acquisition from high-dimensional video inputs for humanoid robots (2015)
    - POWERPLAY (2013)
    - VIME
    - Exemplar Models (EX2)
    - Intrinsic Curiosity Module (Self-supervised prediction)
    - Random network distillation (RND)
- Credit assignment & Sparse rewards
- Training stability
  - Estimator variance reduction
  - Double DQN
  - Dueling DDQN
  - Distributional RL
- Generalization
  - Graph-based
- Scalability
  - Parallelization
  - Sparsity
    - Self-Delimiting Neural Networks
  - Modularization
    - Modular Networks
- Data efficiency
  - Hindsight
    - Hindsight Experience Replay
    - Hindsight Policy Gradients
  - Model based
    - I2A
    - MBMF
    - MBVE
    - STEVE
    - ME-TRPO
    - MB-MPO
    - C & M Architectures
  - Replay
    - Prioritized experience replay
    - Selective replay
    - Distilling / Cloning / Compressing / Collapsing
  - Unsupervised learning
    - Latent representations
    - Hierarchical predictive coding
    - Methods
      - PowerPlay
      - Variational Intrinsic Control (VIC)
      - Diversity is all you need (DIAYN)
      - Variational Option Discovery (VALOR)
    - Auxiliary tasks
    - Minimax training
    - Competition
  - Multitask
    - Universal Value Function Approximators (UVFA)
    - Auxiliary tasks (UNREAL)
    - IU Agent
    - Mutual Alignment Transfer Learning (MATL)
  - Meta-Learning
    - Evolutionary principles in self-referential learning, or on learning how to learn: The meta-meta-... hook. (Schmidhuber Diploma, 1987)
    - On learning how to learn learning strategies. (1994)
    - Reinforcement learning with self-modifying policies. (1997)
    - Gödel machines: Fully Self-Referential Optimal Universal Self-Improvers. (2006)
    - Optimal Ordered Problem Solver. (2004)
    - Steps towards `self-referential' learning. (1992)
    - A self-referential weight matrix. (1993)
    - RL^2
    - Learning to reinforcement learn
    - MAML
    - SNAIL
- Lifelong learning
  - Modularization
    - HRL
    - Progressive Neural Networks
    - PathNet
    - Learning an Embedding Space for Transferable Robot Skills
    - POWERPLAY
  - Memory

http://louiskirsch.com