Question / Idea / Topic / Things I want answers to and want to educate myself on	Related Work
What is the right *general purpose* objective function?	
How useful is the MDP framework for representing (or approximating) general "rational" preference structures in sequential processes, and how can we use it as part of a more general purpose framework?	Axiomatic sequential decision theory from 60s-70s [Koopmans, Kreps, Sobel]; Inverse RL; Preference-based RL; GVFs
How to apply social choice to designing an objective function? How to benefit from multiple teachers when their instructions push agent toward different local minima?	
How to integrate ordinal and cardinal feedback signals in a theoretically-justified way?	Related to social choice topic; Inverse RL, Policy Shaping, Preference-based RL
Working with multiple representations of a thing	
Strength of representations / combining various representations How to represent confidence and uncertainty (is there a difference?) of a prediction? How to merge two estimates of different confidences? How to use two separately trained models (on similar data) to train each other (through bootstrapping)? How do enable useful partial supervision?	
Generating multiple representations via abstraction How to support **sets** of data in a neural architecture / embedding in a theoretically-justified way? Preferably natively, as part of the architecture, so we could take intersections, unions, differences, subsets, etc.	Generative models / latent embeddings; State abstraction in RL [Li Littman Walsh]; multiple representations (can represent sets at different levels)
Good Old-Fashioned AI with Discrete Abstractions How can we use planning/search, logic, frames, etc. to enhance ML-models? How can we translate ML models to symbolic reasoning? Note: There is a feedback loop, which I call "consistency-based learning" **Note**: Ad-hoc justification falls out of this	Includes ML → symbolic reasoning (http://cognition.ouc.ac.cy/learning/) but also reverse: symbolic reasoning → ML, which I haven't seen.
Generalization	
How to characterize the "locality" of learning: e.g., learning something specific, while still generalizing, without displacing other knowledge in different areas of the data space [cf. Catastrophic forgetting, but not necessarily on separate tasks/domains] [Maybe this is solved practically via modular architectures, and consistency-based learning; but would still be useful to quantify it in some way] [Cf: 1-step TD has extremely low variance and eventually outperforms multi-step TD lambda / source traces because of their high variance (they cause forgetting in other parts of the state space); Cf. Gordon 1995 work function appx in RL]	