Consensus (N = 2) MDP loss, property steps_max, shallow 3-layered network

- One-hot encoding
- Ordinal encoding
- One-hot encoding (Q-table)
- Ordinal encoding (Q-table)
- One-hot encoding (Double Q-learning)
- Ordinal encoding (Double Q-learning)
- One-hot encoding (Double Q-learning, Q-table)
- Ordinal encoding (Double Q-learning, Q-table)