overlooked multigent environment

Starcraft multi agent challenge

Swarm particle simulation

Curriculum adversarial training

Ensemble RL

Nanocloud ancient Egypt domain

resume / website update

blurbs on topics w/ sources

Finish RM2 rate transfer

ME531 Project: simulation: matrix transforms
Flares Env

Paper parser

masters rough draft

auto curricula proposal

more support types / methods of support

support as an action

Population migration

Fetero/RM2 integration

Masters registration

credits, committee, paperwork

QD in human diversity proposal

Vision state representation

Survey papers

Fitness modeling

Env generation

QD methods

Island models

Frame reliance as a tradeoff of individualized skills
and all learning simultaneously

relevance to field/contribution
cannot have everyone possibly switch behaviors

MTL using own behaviors versus relying on others
shared policies all learn same thing

tell other agents
how to act/
what to do

QD function on island
idea on mainland
Another island multiobjective

islands produce initializations
for optimization process
running on mainland
EA's prone to local minima
based on starting conditions

Q1 - look at diversity as an inherent characteristic based on the network, not the behavior
- i.e. behavior might not be known since the changes are careful and specific in order to achieve a desired behavior
- i.e. the alteration may or may not have a direct impact on the expressed behavior and the reachable set is based on the weights altered
- if Q1 is the 1st step, behaviors are what we are looking for, but if it is a step before optimization, then we would care about diversity in how they can be altered
- basing diversity on characteristics of the behavior is also very domain specific and determining it to determine well and robustly
- from bounded reachability analysis

[illegible] review   Bias <-> var

weight ^simple policies based on state features

Non stationarity is a core challenge in multiagent learning. Some methods of dealing with it break learning into several phases, where the first phase learns individual skills, and then a separate optimizer learns to select between these higher level skills, effectively reducing the requirement that agents stumble upon random joint actions that lead to an eventual reward. (SMFL, maddpg, options)

Another approach employed by CLEAN rewards is to freeze the policies of other agents in the system and cycle through the agents such that only a single agent is actually updating its policy at any given iteration. The challenge here is that this greatly increases the computational burden of the learning process, and makes the process more likely to converge to a local optima due to additional factors such as the ordering the agents are cycled through or how many iterations each agent is allowed to learn before freezing it

and allowing another agent to update its policy.
In this paper

---

individualize[?] otherwise[?]
Bring up individualistic behaviors when discussing weaknesses of multi-stage approach

---

program of study

committee Davidson
Gratt[?]

Pull from second
question in Quals
defend week of 11th
not 2-4
13 14 11 12
6 7 8 } fallback
