Purpose:

Monoculture is a problem in plants. If many plants are genetically identical, a virus can easily wipe them out. We believe this problem also exists in robots. If, for example, all autonomous cars run on the same software, and someone finds a way to exploit that software, they can easily cause mass destruction. It is therefore better to have diversity in the cars, such that if one is exploited the others are not.

Basically, in a training environment diversity is preferable to superior performance, when looking for maximum performance when exposed to new environments.

Method:

In order to prove that this is the case, we want to show that a group of diverse robots, when faced with a new and potentially hazardous obstacles, will outperform a group of homogenous robots.

* We will create three distinct groups of individual robots, called fleets.
* Fleet A contains a monoculture (all robots are identical), fleet B contains robots with distinct neural networks, and fleet C contains robots with distinct neural networks and morphologies.
* Each fleet will contain 10 ‘champions.’
* A champion is the most successful individual that emerges from a parallel hill climber.
* In the case of fleet A, 10 champions will be selected, but only the best (most fit to training environment), will be kept. It will just be copied 10 times.
* For the other two fleets all champions will be kept.
* We then expose these fleets to a new environment and observe their performance.
* We get the fitness of the one robot from fleet A, and of the best robot from fleet B (or C).
* This is repeated 30 times (more if necessary).
* The collection of 30 fitness values from either fleet is called a “squadron.” 30 values originally from fleet A create squadron A.

Null hypothesis: It is equally likely that the fitness of a randomly selected individual from one squadron will be higher or lower than the fitness of a randomly selected individual from the other squadron.

Alternate Hypothesis: It is more likely that a randomly selected robot from “squadron A” has a lower fitness than randomly selected robot from “squadron B”

We will rank and compare the 30 data points from squadron A with the 30 data points from squadron B using a Mann-Whitney U test.

This process is repeated to compare squadron A with squadron C and squadron B with squadron C.

We’re saying that you have a task which you plan to solve with a robot. You develop a training environment as close as possible to the eventual task, but it can’t be perfect. You are better off with a fleet of diverse robots, then with one that is well tuned to the task.

Im just struggling here with if this really says anything.

What the actual hypothesis? Is it just that