Modeling Changes in Exploitative vs. Protective Behavior

Joseph Blass

Motivation and Questions

- Humans exploit others for selfish reasons
- Humans also protect other people against exploitation
- These behaviors often more or less balance each other out
- What circumstances in society incentivize its members to exploit each other, protect each other, or leave each other alone?
- Are these traits that can be selected for in a population?
- Can we build a model that selects for such traits and balances them against each other?

I built a NetLogo model, and a HubNet model to explore this question space collaboratively

Background

- Findings from psychology [1] and biology [2] support the idea that certain psychological traits, such as a tendency to cooperate and to punish non-cooperators, are innate
- “The first rule of life in a dense web of gossip is: Be careful what you do. The second rule is: What you do matters less than what people think you did.” [3]
- Game theory models of evolutionary ethics show that self-interested agents can learn to cooperate and develop heuristics like the golden rule [4]

The Model: Parameters

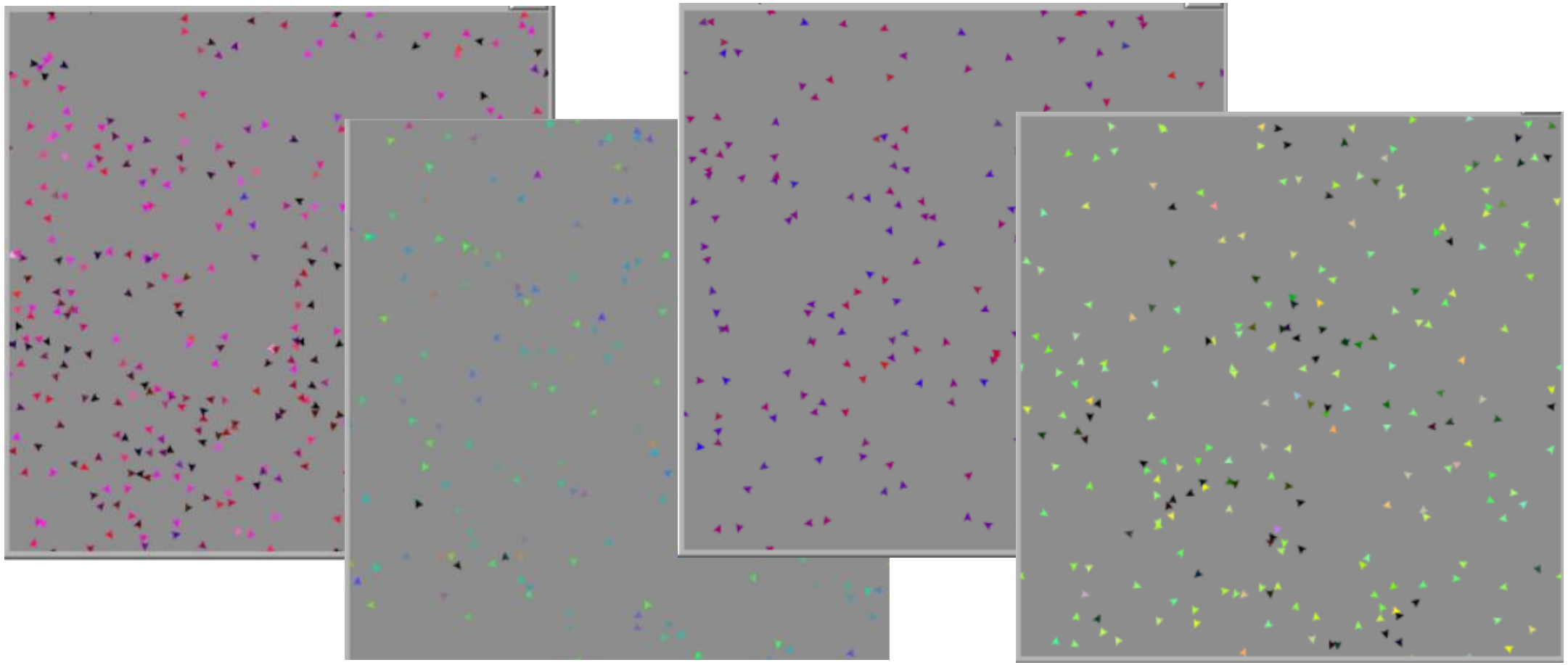
- Social currency is a stand-in for wealth, status, and reputation. Every turtle starts with the same amount of social currency
- Each turtle has a likelihood of stealing, protecting, or doing nothing. During set-up, these variables are initialized to user-set parameters (plus some randomness)
- ‘Probability of getting away’ determines how likely turtles who steal are to avoid getting caught
- Turtles above a social currency threshold can reproduce. Offspring have the average of their parents’ parameters, plus some randomness
- Turtles die if they reach life expectancy or have no social currency

The Model: Actions

- For all turtles, $0 \leq$ Theft-threshold $<$ Protect-threshold ≤ 100
- Each turtle picks a number X, where $0 \leq X \leq 100$.
  - If $X <$ Theft-threshold, the turtle steals. Thieves pick a second random number to determine if they get caught
  - If Theft-threshold $\leq X <$ Protect-threshold, the turtle does nothing
  - Otherwise the turtle protects
- Thieves who get away or are not seen by protectors siphon social currency from their neighbors
- Protectors who catch a thief in their neighborhood get social currency from that thief. Otherwise they give up social currency to their neighbors as punishment for being nosy busybodies
- Turtles who do nothing gain or lose social currency based on their neighbors' actions

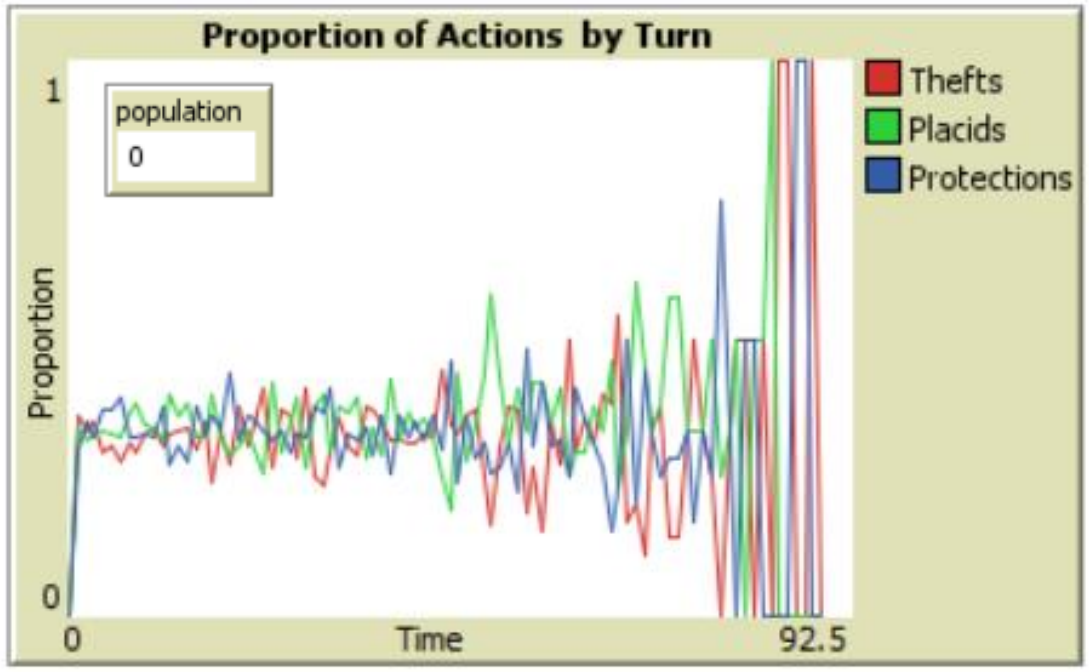
Color and Visualization

- Turtles were color coded for visualization purposes: thieves turn redder; protectors turn bluer; do-nothings turn greener
  - Background is grey so we can see white (mixed) and black (new) turtles



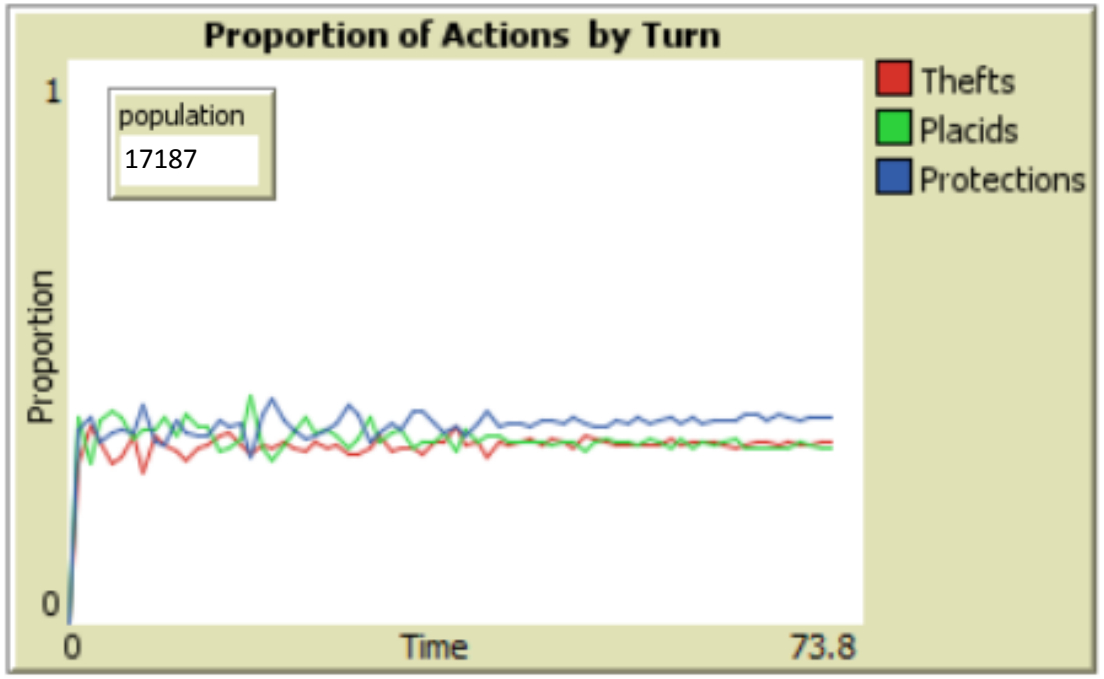
Results

- Many settings were unstable and quickly led to population collapse:



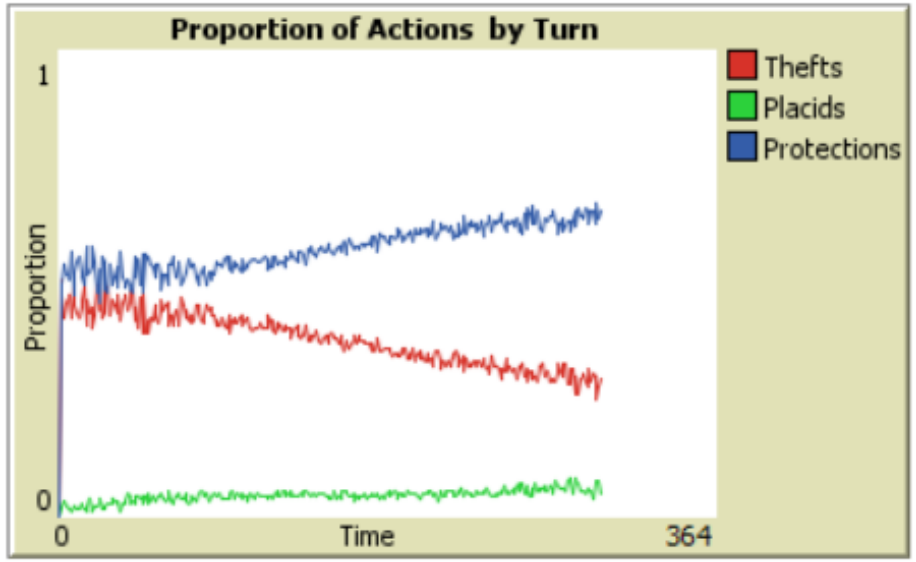
Run lasted 88 ticks with theft, protection, & doing nothing equally likely; probability of getting away = 35; initial social currency = 20; and mate threshold = 80.

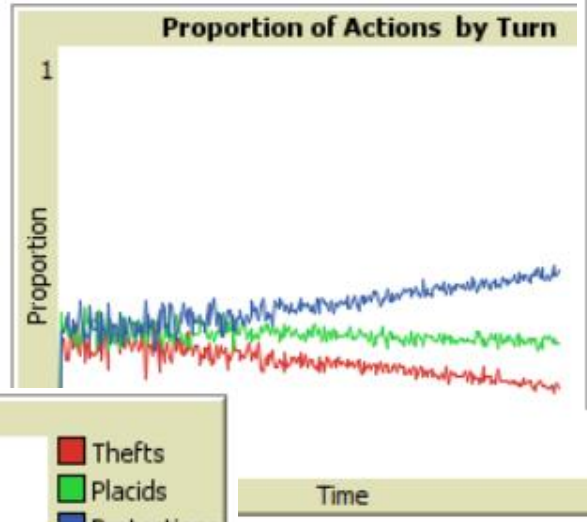
- If the barrier to mating was low, population exploded and traits were not selected for:

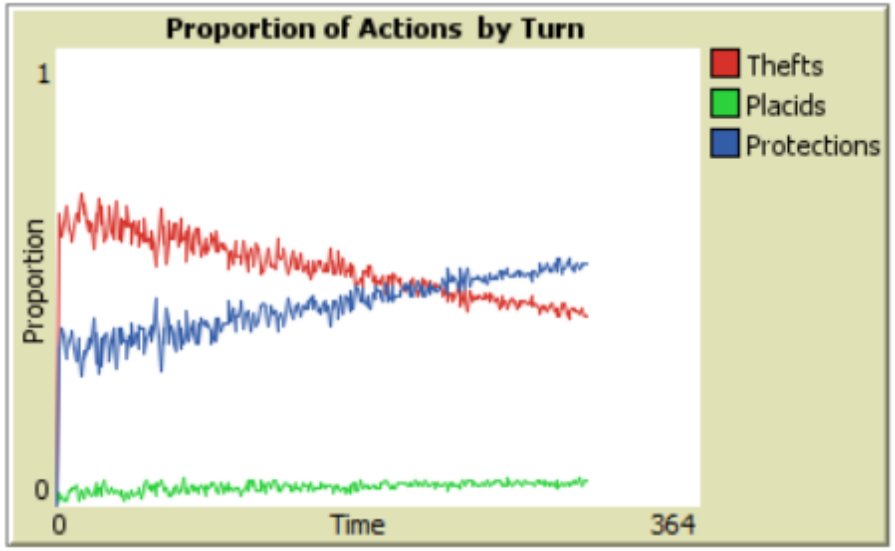


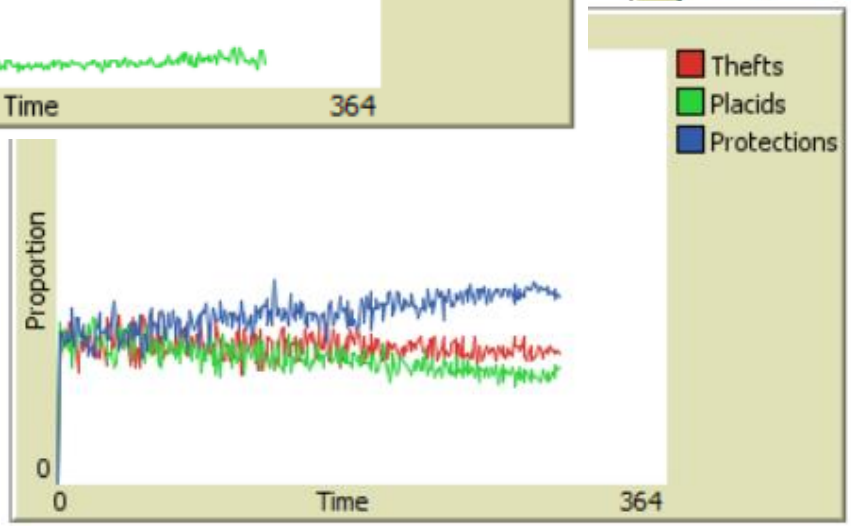
Same settings as on the left, but with mate threshold = 40. After 71 ticks, population = 17,187. Protection is very weakly selected for.

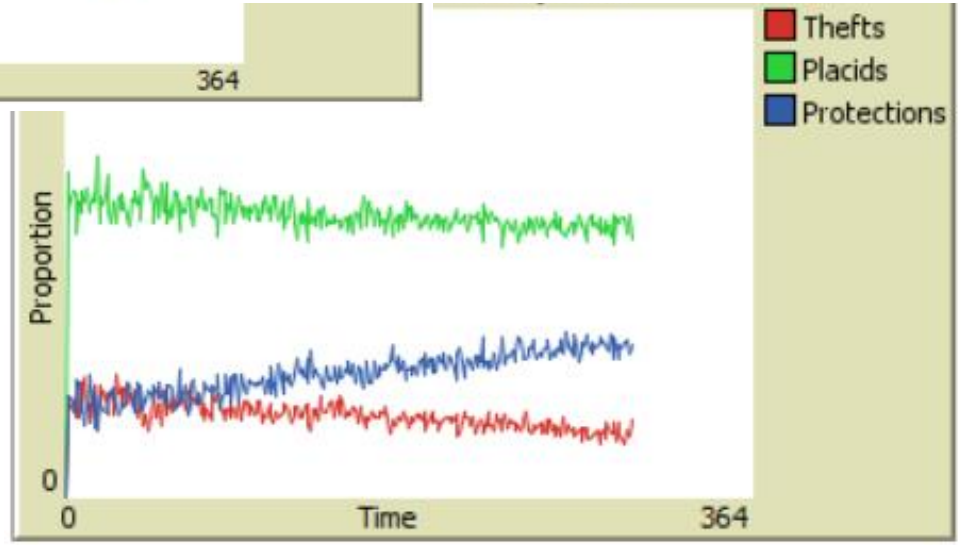
- An action is selected for if it is taken more frequently over time
- Most settings that did not result in population extinction or explosion selected for protecting as a trait





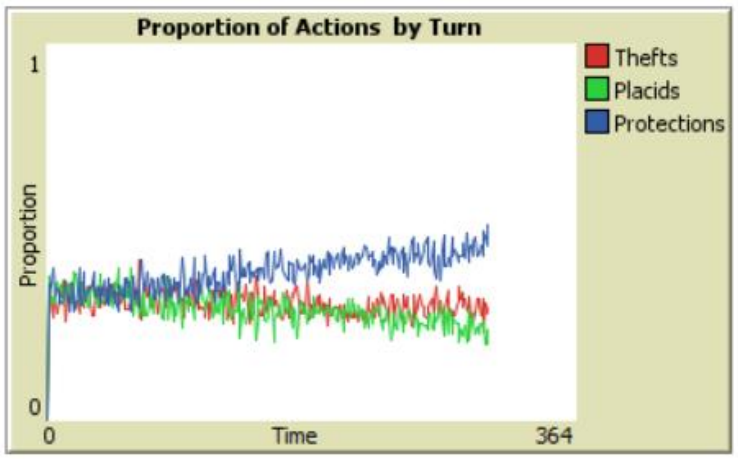




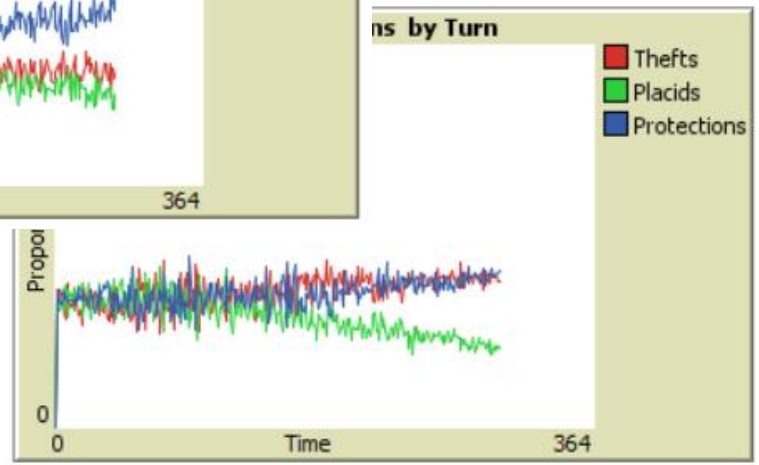


A selection of different parameter settings run to 300 ticks.

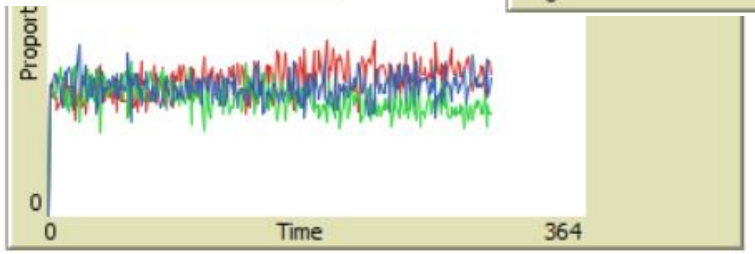
- If probability of getting away was high enough, theft was selected for



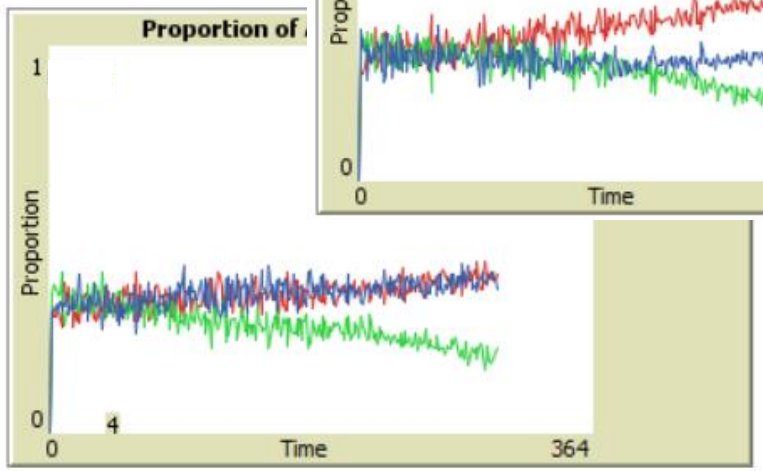
p-getaway = 40



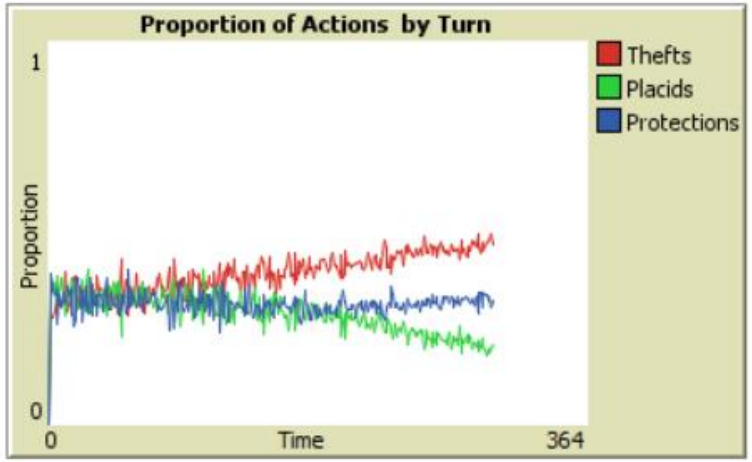
p-getaway = 50



p-getaway = 60



p-getaway = 70



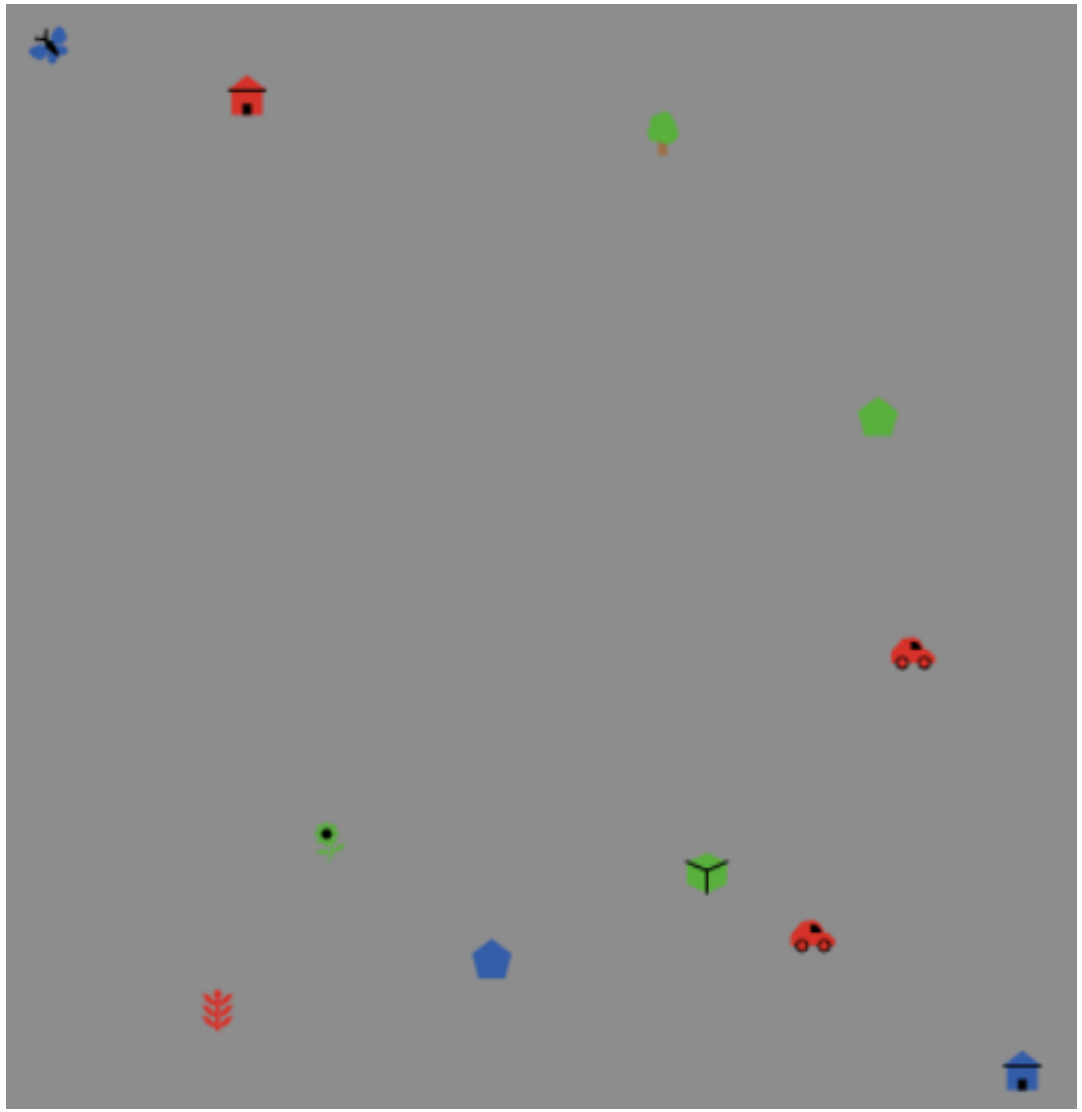
p-getaway = 80

Several runs to 300 ticks. Equal initial probability of theft, protection, and doing nothing. The only change from run to run is increasing the probability of getting away.

HubNet (Multi-User) Model: Simplified Behavior

1. People caught stealing last turn move first; then all others move
2. All people select actions
3. Probability of getting away decided using randomness
4. People caught stealing turn red; protectors turn blue; others turn green.

- No death or reproduction.

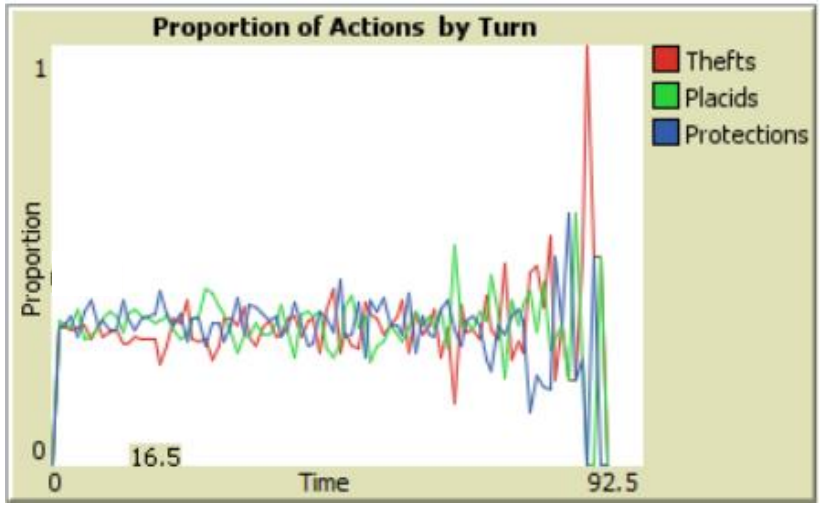


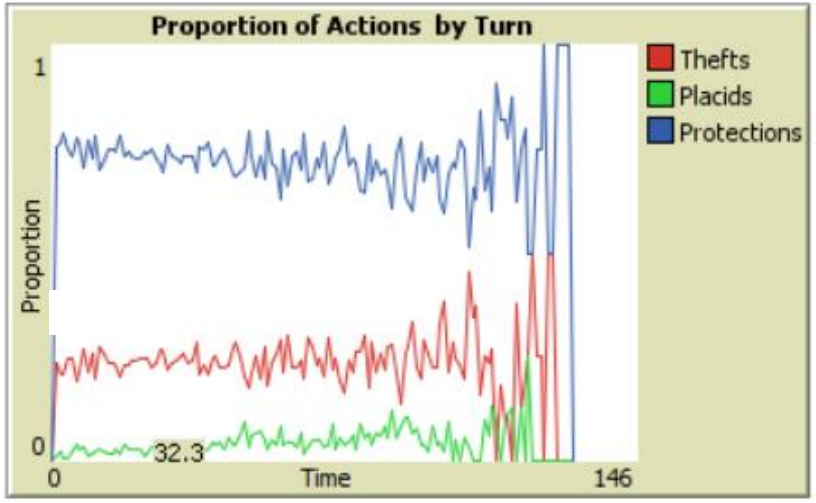
Conclusions

- The model was able to demonstrate population change and selection of traits over time
- The model usually selected for protection, but would select for theft if the likelihood of getting away was sufficiently high
- Several critical points were discovered that changed which traits were selected for, or whether the model would run at all
- I was unable to find parameter settings that establish an equilibrium within the model. If the population did not die out or explode, one trait was always selected for
  - However there were parameter settings that changed between traits selected for from run to run

Limitations

- Not a true model of altruism: protectors had a direct incentive to protect, since they were also protecting themselves
- Not enough incentive to do nothing
- Individual turtles should have individual likelihoods of getting away with theft or catching a thief when protecting
- Too much reward for protecting: unless getting away was very likely, protectors were likely to find a thief





Future Work

- Expand cognition of turtles
  - Have them take into account the probability of getting away
  - Have them consider the past actions of the turtles around them
- More individualized parameters, including probability of getting away and probability of catching thieves

Selected References

[1] Haidt, J., & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, *133*(4), 55-66.

[2] West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Current Biology*, *17*(16), R661-R672.

[3] Haidt, J. (2007). The new synthesis in moral psychology. *Science*, 316(5827), 998-1002.

[4] Skyrms, B. (1996). *Evolution of the social contract*. Cambridge University Press.