826G5Z - Decision Making in Ants

Candidate No. 269648

1. Methods: Describe the algorithm

The ant decision making model is based on Robinson et al.'s work [1] in 2011, which uses a simple threshold rule to explain the ant collective decision-making. In 826G5Z module, the standard methods and parameters are contributed by Chris Johnson, Andy Philippides, and Alex Dewar.

In the paper [1], the home site is considered uninhabitable (in python can be set as -infinite), and ants can switch between assessment states. The model assumes ants continue searching until they encounter a nest of sufficient quality, and may revisit the same nest based on a certain probability, but do not remember previous visits. when the site's quality is higher than an ant's individual threshold, the ant commits to a site, it remains committed and recruits nest-mates, therefore, the commitment states are absorbing states in the Markov process.

1.1 Robinson.py

The main methods are written in robinson.py, the model there uses Monte Carlo simulation [2] which repeats sampling to obtain the statistical properties of ant colony's decision process, so the number of samples is the number of ants in the colony.

1. plot_ant_path: This function visualizes the path of an individual ant.

2. plot_summary_data: This function is designed for plotting summary data and aggregating results from the ant colony. It can represent the distribution of sites ultimately selected by the ants, the time distribution of ants making final decisions, the mean time each site was discovered, and the mean number of visits per site.

3. plot_summary_data_show_num_selected: This is a variation of the plot_summary_data function, it includes additional details the number of ants selected the new nest and the total number of ants (colony size) in the plot.

4. run_model: This function architects the main model for simulation, and it invokes the plot_ant_path functions to visualize the individual ant path. The function simulates each ant's interaction with various sites, it inputs the parameters set in lab4_robinson.py (explain in the following texts), and output the times to first recruitment, site first discovery times, site visits times, accepted sites, random seed, and the attributes of each ant (including path, the time when ants visit the sites on the path, threshold, final selected site, number of steps taken).

The use of standard deviations, means, and the random number with normal distribution leads the model involves probabilistic elements and could be simulating random variables (for example ant travel times and quality assessments).

For the code skills, the function uses two loops. The ‘for i in range(n):’ block lets each ant in the colony to complete a simulation of site decision making. And the ‘while True:’ block lets ants continue search the site until the site perceived quality higher than the threshold or the travel steps exceed the maximum number of steps.

1.2 lab4_robinson.py

And in lab4_robinson.py, the codes first defined many variables and assigned values to them, then invoke the run_model function of robinson.py and check the summary by invoking plot_summary_data function. Additionally, after simulation, the ants state information are saved in a pkl file for easier check and model analysis. The experimenter can adjust the following parameters and run the file to view simulation results:

1. time_means: mean time to get between each nest, use a 2d array, the column means the start site, and the row means the target site, so the numerical value represents the distance between two sites.
2. time_stddevs: standard deviation of time to get between each nest, because the movement speed of the ants is influenced by randomness in real world.
3. probs: probabilities of ant visiting each nest from each nest, the data structure is also a matrix.
4. n: the number of ants (colony size).
5. threshold_mean: the mean threshold of ants, it will be compared to the quantity of sites.
6. threshold_stddev: the standard deviation of ants threshold, because the cognitive situation of each individual ant is a little different.
7. quals: a 1d array of the mean quality of each nest, it is the ants’ perceived quality of each nest essentially.
8. qual_stddev: standard deviation of nest qualities, it means how variable the ants assessment of each nest is, because the ants assessment can be influenced by their physical strength and environment.

2. Run the model with standard parameters

When running the model with standard parameters in figure 1, the summary results (figure 3) and the individual runs results (figure 4) are visualized. Cause compared to function plot_summary_data, function plot_summary_data_show_num_selected can additionally display the number of ants which have selected a new nest and the ant colony size, the plot_summary_data_show_num_selected function is used for analyzing the summary results (figure 2).

```python
# mean time to get between each nest 每个巢穴之间的平均时间
time_means = np.array([[1, 36, 143], [36, 1, 116], [143, 116, 1]])
# standard deviation of time to get between each nest 每个巢之间到达时间的标准差
time_stddevs = time_means / 5
# probabilities of visiting each site from each other 从此访问每个站点的概率
probs = np.array([[0.91, 0.15, 0.03], [0.06, 0.8, 0.06], [0.03, 0.05, 0.91]])
# set the number of ants 蚂蚁数量
n = 27

# these govern the ant's threshold 蚂蚁阈值
threshold_mean = 5
# threshold_mean = 2.5
threshold_stddev = 1
# mean quality of each nest. Note home is -infinity so it never gets picked
# 每个巢的平均质量
quals = np.array([-np.inf, 4, 6])
# standard deviation of quality: essentially this controls
# how variable the ants assessment of each nest is. This is currently set
# as in the 1st experiment where the variability is the same for each nest
# 质量标准差（蚂蚁评估巢穴质量的可变性）
qual_stddev = np.array([1, 1, 1])
# However, if you want to change is so nests perceived w different accuracy
# you could do eg qual_stddev = [1, 1, 4]  巢穴的感知精度
```

Figure 1. The standard parameters

```python
state = robinson.run_model(n, quals, probs, threshold_mean, threshold_stddev,
                           qual_stddev, time_means, time_stddevs, [])

# Save the data for this run of the experiment
with open(output_file, 'wb') as file:
    pickle.dump(state, file)

# NB: To read the data again in another file, you can do e.g.:
# with open('robinson_data1.pkl', 'rb') as file:
#     state = pickle.load(file)

# Unpack the tuple...
current_time, discovers, visits, accepts, Ants, rnd_seed = state
# 蚂蚁首次招募之前的时间（时长），每个站点首次被蚁群发现的时间，每个站点被蚁群访问的次数，每个蚂蚁接受的站点，每个蚂蚁的属

# Plot Summary data
# robinson.plot_summary_data(current_time, accepts, discovers, visits)
robinson.plot_summary_data_show_num_selected(current_time, accepts, discovers, visits, Ants)
```

Figure 2. Function plot_summary_data_num_selected is invoked in lab4_robinson.py

In figure 3, the plots show the ant path changing with the steps and times (x axis), the y axis means the sites, site 0 is the home site, and the site 1 and site 2 are the new assessment sites, so the red lines denote the path of the ant moving between different sites. The chart title prompts the ant threshold, the perceived quality of the selected site, and the number of steps and time taken during the travel process.

Among the three ant paths from left to right:

1. The large search costs 129 steps with about 722 times, but the little search costs 3 steps with about 149 times, it can be evidenced that once the site perceived quality reach the ant threshold, the ant immediately selects the site regardless of the time (searching time is short or long), steps (searching steps are many or little), or and the number of searched sites (ant searched for a new site or two new sites).

2. Due to the non 100% probability of ants visiting a new site, ants may spend multiple steps at the same site.

3. During the search process, after an ant goes to a new site, it is still possible for it to return to the old site.

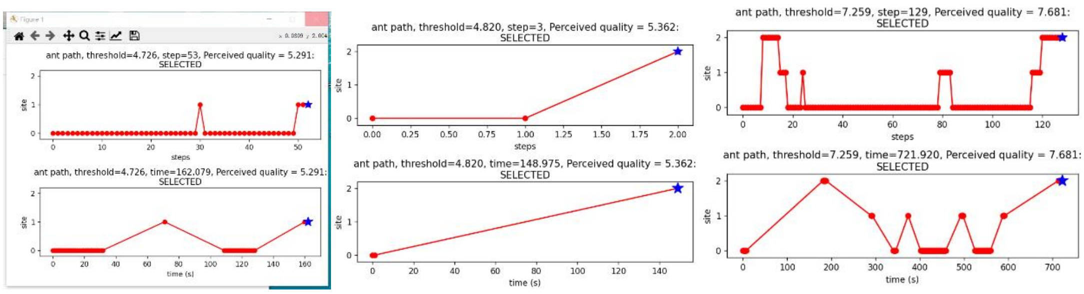

Figure 3. The individual results with standard parameters

The figure 4 shows two histogram charts and two bar charts, they represent the summary results of the ant colony:

Top-left chart shows the distribution of the final site selection by the ants. It titles "27/27 selected a site", and the number of ants in site 0 (home) is 0, it indicates that in the condition of standard parameters, all ants can select the new nest. Additionally, sites 2 is chosen more frequently than site 1, because the site 1 mean quality (4) < ant mean threshold (5) < site 2 mean quality (6), the site 2 perceived quality can satisfy more ant thresholds.

Top-right chart indicates the time it took for ants to make their final decision. Most ants made their decisions quickly (under 300 seconds), with fewer ants taking longer.

Bottom-left graph displays how long in average it took ants to discover each site. Site 0 (home) was discovered the fastest (ant stay in home at the beginning), and site 2 took the longest time to be discovered, the reason could be: site 2 is farther from home than Site 1; and site 1 can't reach some ants' threshold, when these ant visit the site 1, they would not select site 1 and need to move and selected the higher perceived quality site (like site 2).

Bottom-right chart shows the mean number of visits to each site, it shows home > site 1 > site 2. But for the final site selection, site 2 > site 1 > home. This suggests that in this model, ants do not choose a site based on familiarity by visiting it multiple times.

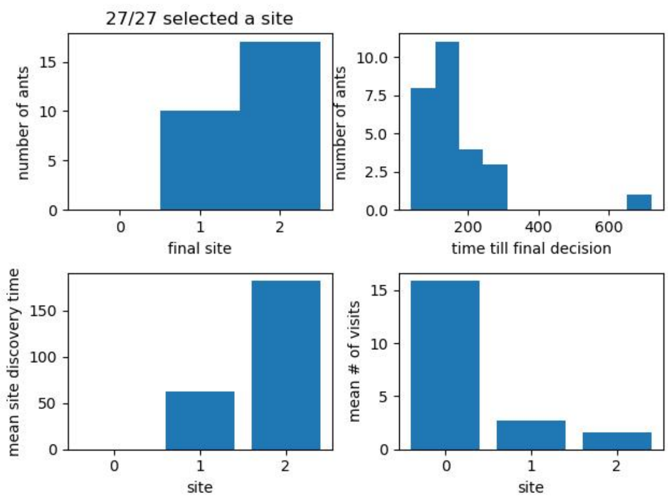

Figure 4. The summary results with standard parameters

3. Change the threshold mean

In order to avoid the incidents, the experiments of Q3-Q5 are repeated five times for each setting of threshold mean, therefore, this question will be analyzed using charts that represent the typical situation of five sets of experiments for each threshold mean.

3a. Halved threshold mean

When use halved threshold mean (2.5), the summary results are shown in figure 5.

1. All ants can select the new nest, it is same to the standard threshold mean (10), because mean threshold (2.5 or 5) is less than the mean site 2 quality (6).

2. Different from the standard situation, in the final decision, number of ants in site 1 is larger than site 2. It is because the mean quality of site 1 and site 2 all satisfy the mean threshold, but the site 1 is closer to the home, ants are easier to arrive site 1.

3. The mean number of visits for each site is reduced, and the time till final decision distributes under 250 time. The reason is that the closer-distance site 1 quality can satisfy the halved threshold mean, which avoid the situation that ants first find site1 and then search for a new and better site due to not meeting the threshold. Therefore, the ant path become short (figure 6), the time and steps become small.

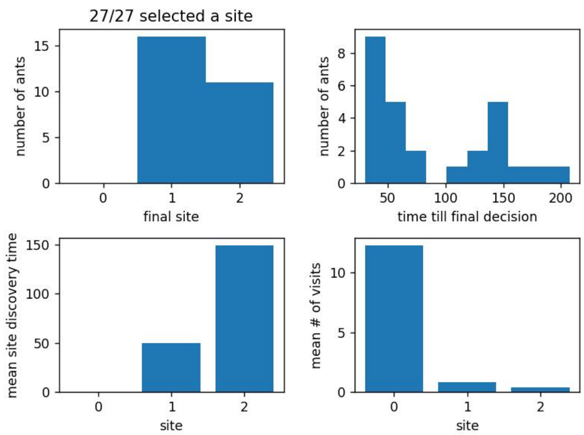

Figure 5. Summary results with halved threshold mean

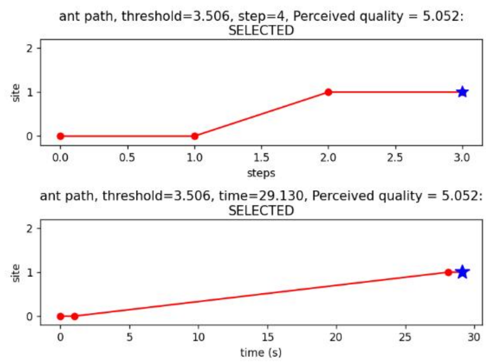

Figure 6. Individual results with halved threshold

3b. Doubled threshold mean

Doubled threshold mean is 10, the summary results and individual results are shown in figure 7 and figure 8.

1. Only 5/27 ants selected a site; most ants haven't selected a site until reach the maximum steps limit. Also, the mean number of visits for each site is increased, and the time till final decision can even distributed above 10000. Due to the doubled threshold mean (10) > site 2 mean quality (6) > site 1 mean quality (4), most ants never finding a satisfactory site, will complete the maximum step (1000), which increases the search process time and the number of visits to each site.

2. The mean site discovery time for each site is similar to the standard or halved threshold mean, because the nest situation (time to get between each nest and the probabilities of visiting each site from each other) is not changed.

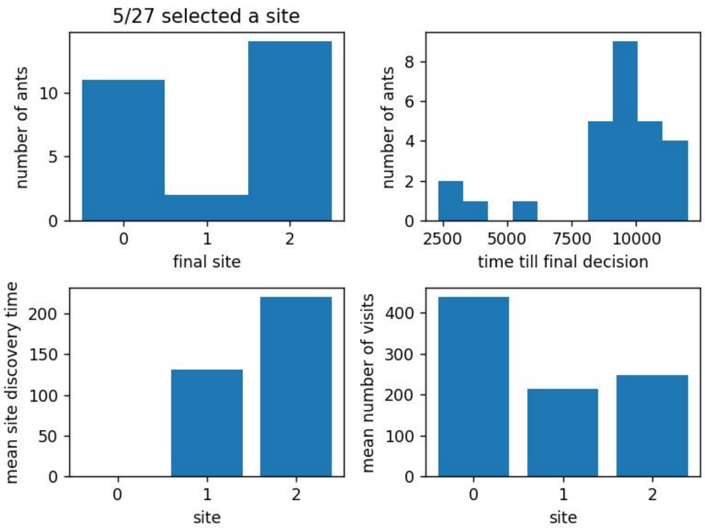

Figure 7. Summary results with doubled threshold mean

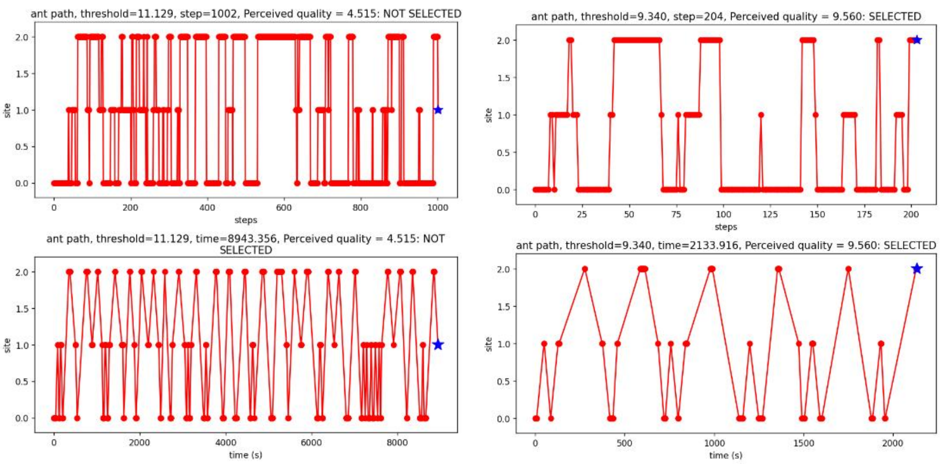

Figure 8. Individual results with doubled threshold

3c. Change the threshold standard deviation

The halved threshold standard deviation (value is 0.5, figure 9 and 10) and the tripled threshold standard deviation (value is 3, figure 11 and 12) are experimented. The findings are:

The smaller threshold standard deviation makes the ants' threshold close to the mean threshold (5), so distribution of the time till final decision are concentrated.

The bigger threshold standard deviation makes the distribution of the time till final decision be very scattered, some ants selected the site quickly (about less than 100 time), but some ants even can't select a satisfied site in the limited steps. This is due to the big threshold standard deviation leading to the very small threshold (threshold = 0.783 in the right of fig 12) or the very large threshold (threshold=9.174 in the left of fig 12), the probability of reaching a perceived quality of 9.174 is very low for new sites with an average quality of 4 or 6.

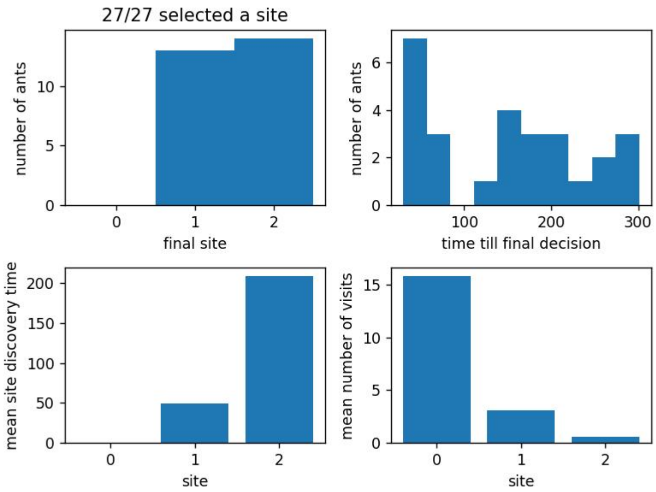

Figure 9. Summary results with halved threshold standard deviation

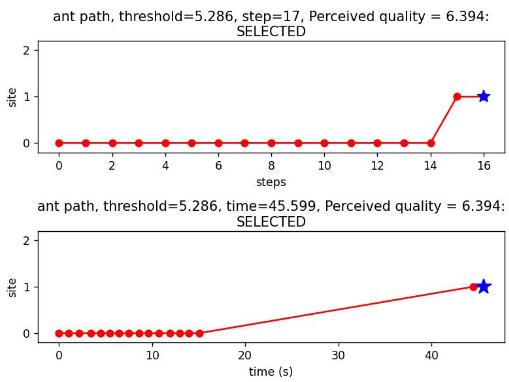

Figure 10. Individual results with halved threshold standard deviation

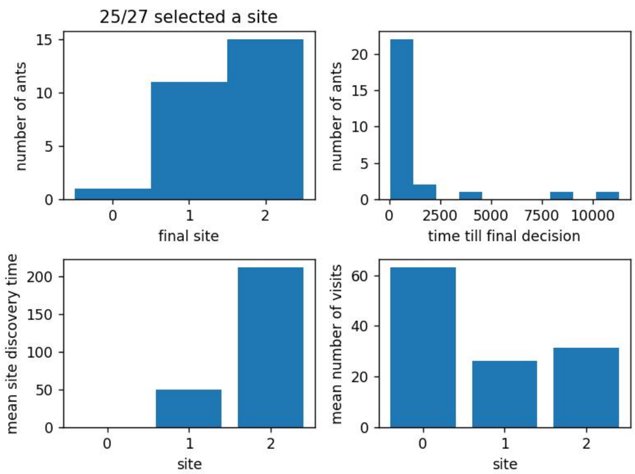

Figure 11. Summary results with tripled threshold standard deviation

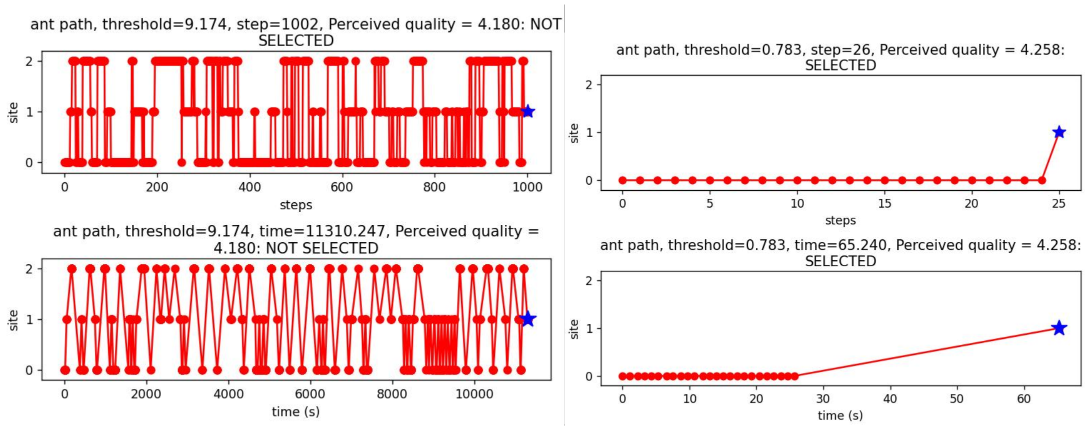

Figure 12. Individual results with tripled threshold standard deviation

4. Change the symmetric nest situation

In this part, the nest situation is changed to be symmetric, the nest map (Figure 13) and the nest parameters (Table 1) are follow the Robison et al.'s paper [1].

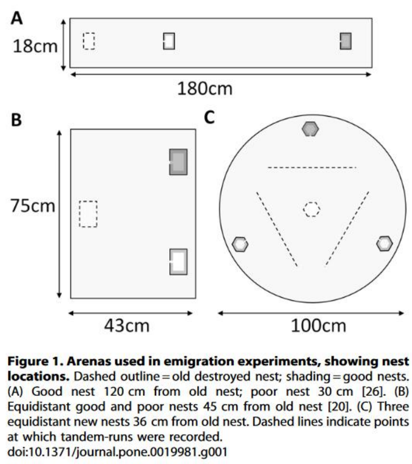

Figure 13. Nest map of the symmetric nest situations

Table 1. Nest parameters of the symmetric nest situations

Table 1. Parameterisation used in simulations of Monte-Carlo model.

Parameter	Comparison with [26]	Comparison with [20]	Comparison with new multiple-nest experiment	Derivation
Number of nests	3	3	4	From experiments
Arena size and shape	See Fig. 1a	See Fig. 1b	See Fig. 1c	From experiments
Position of nests	Good nest (A) further than poor nest (B) (Fig. 1a)	New nests equidistant from old (Fig. 1b)	New nests equidistant from old (Fig. 1c)	From experiments
Mean travel time between nests (sec) from column nest to row nest (SD = 1/5 mean)	$\left(\begin{array}{cccc} & \text{Old} & \text{A} & \text{B} \\ \text{Old} & 1 & 36 & 143 \\ \text{A} & 36 & 1 & 116 \\ \text{B} & 143 & 116 & 1 \end{array}\right)$	$\left(\begin{array}{cccc} & \text{Old} & \text{A} & \text{B} \\ \text{Old} & 1 & 54 & 54 \\ \text{A} & 54 & 1 & 81 \\ \text{B} & 54 & 81 & 1 \end{array}\right)$	$\left(\begin{array}{ccccc} & Old & A & B & C \\ Old & 1 & 46 & 46 & 46 \\ A & 46 & 1 & 80 & 80 \\ B & 46 & 80 & 1 & 80 \\ C & 46 & 80 & 80 & 1 \end{array}\right)$	From walking speed 8.4 mm/s [74]
Probabilities of finding nests (from column to row)[1]	$\left(\begin{array}{cccc} & \text{Old} & \text{A} & \text{B} \\ \text{Old} & 0.91 & 0.15 & 0.03 \\ \text{A} & 0.06 & 0.80 & 0.06 \\ \text{B} & 0.03 & 0.05 & 0.91 \end{array}\right)$	$\left(\begin{array}{cccc} & \text{Old} & \text{A} & \text{B} \\ \text{Old} & 0.70 & 0.15 & 0.15 \\ \text{A} & 0.15 & 0.70 & 0.15 \\ \text{B} & 0.15 & 0.15 & 0.70 \end{array}\right)$	$\left(\begin{array}{ccccc} & Old & A & B & C \\ Old & 0.76 & 0.08 & 0.08 & 0.08 \\ A & 0.08 & 0.82 & 0.05 & 0.05 \\ B & 0.08 & 0.05 & 0.82 & 0.05 \\ C & 0.08 & 0.05 & 0.05 & 0.82 \end{array}\right)$	From arena size, arena shape & nest positions; see Text S1.
Number of ants	27 (Fig. 3); 49 (Fig. 4)	29 (Table 2); 12–63 (Fig. 5)[2]	13 (test 1); 20 (test 2)	From experiments
Acceptance threshold distribution (A)[3]	Normal distribution: mean = 5, SD = 1	Normal distribution: mean = 5, SD = 1	Normal distribution: mean = 5, SD = 1	Arbitrary
Nest qualities (b)[3]	Old = −inf; Poor = 4; Good = 6	Old = −inf; Poor = 4; Good = 6	Old = −inf; Poor = 4; Poor = 4; Good = 6	Arbitrary
Assessment error (from which ε is drawn)[3]	Normal distribution: mean = 0, SD = 1	Normal distribution: mean = 0, SD = 1	Normal distribution: mean = 0, SD = 1	Arbitrary

[1]This includes *r*, the probability of re-discovering the same nest.
[2]Numbers of ants are colony specific for recruitment latency simulations; see Fig. 5.
[3]These acceptance threshold distributions and error rates correspond to quality-dependent nest acceptance probabilities of 0.76 for the good nest and 0.24 for the poor nest. See Text S1 for details.
doi:10.1371/journal.pone.0019981.t001

When the nest situation is same as the Figure 13 B, there are equidistant good (site 2) and poor (site 1) nests 45 cm from old nest. Compared to the standard parameters in Q2, the results' changes (Fig 14, 15) are:

1. The mean site first discovery times of good and poor site are similar, it is because the site distances and reach probability are same.

2. Mean number of visits for the site 0 (home) is reduced, because the probability of moving to site 2 has increased (changed from around 0.05 to 0.15) and the probability of moving to site 1 has not changed, therefore the probability of ants stay in home has reduced.

3. The decision-making process has been shortened in time, which can be seen in the distribution of time till final decision in Fig 14. The reason is also that the probability of moving to site 2 has increased.

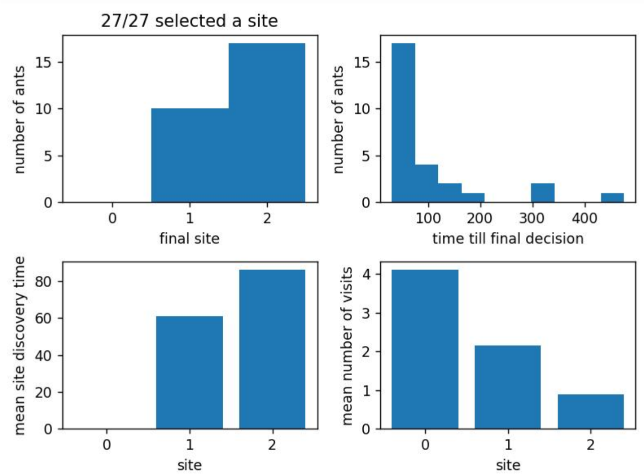

Figure 14. The summary results with equidistant good and poor nests

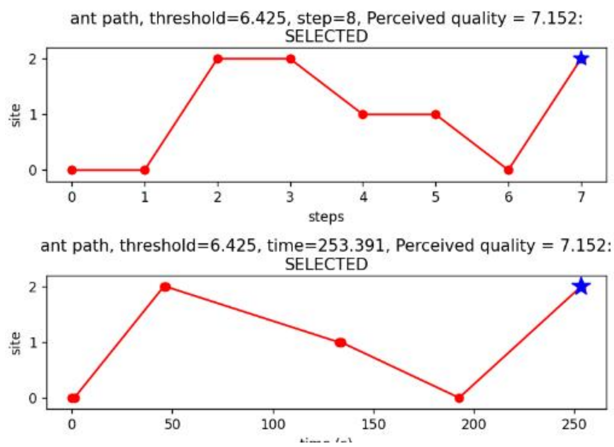

Figure 15. The individual results with equidistant good and poor nests

5. Change the quality standard deviation

The experiment changes the nest quality standard deviation from 1 to 2 (doubled quality standard deviation) for each site. Compared to the standard parameter in Q2, there are two differences:

1. The time till final decision has been shortened, the longest time is around 300, it’s much shorter than the 650 with standard parameters. It is because the bigger quality standard deviation, the bigger probability of the high perceived quality site, and the easier to satisfy the ants with high threshold.

2. For the final selected site, site 1 is chosen more frequently than site 2, this result is opposite to Q2. Due to with standard parameters, poor site 1 can not reach some ants’ threshold; but with high quality standard deviation, the poor site 1 can have higher perceived quality, hence, it can be selected by more ants now.

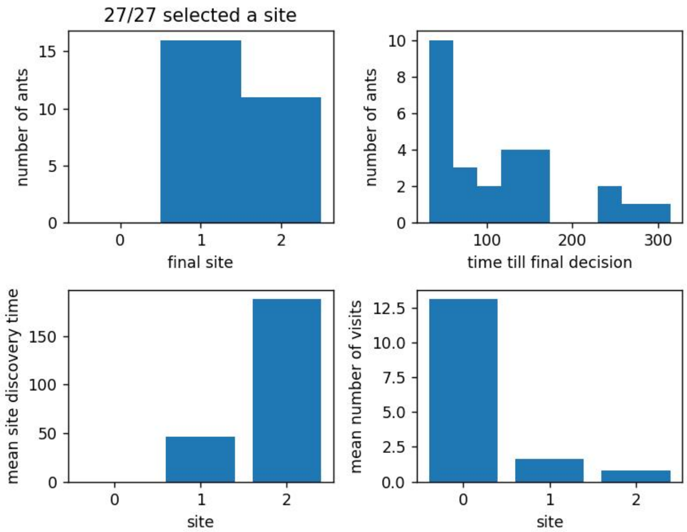

Figure 16. The summary results with doubled quality standard deviation

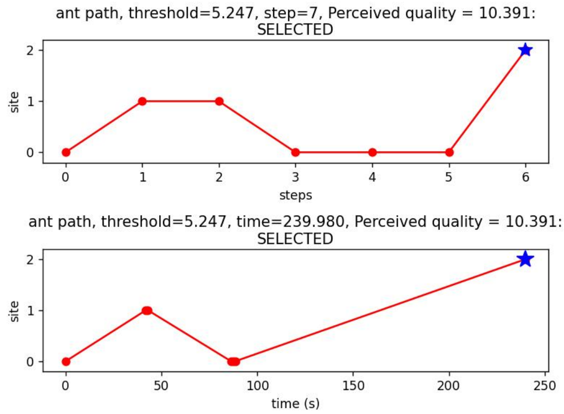

Figure 17. The individual results with doubled quality standard deviation

6. Parameter sweep

The experiments choose the threshold standard deviation to do the parameter sweep, the detailed values are: [0.1 0.3 0.5 0.7 0.9 1.1 1.3 1.5 1.7 1.9 2.1 2.3 2.5 2.7 2.9 3.1 3.3 3.5 3.7 3.9 4.1 4.3 4.5 4.7 4.9 5.1 5.3 5.5 5.7 5.9 6.1 6.3 6.5 6.7 6.9 7.1 7.3 7.5 7.7 7.9 8.1 8.3 8.5 8.7 8.9 9.1 9.3 9.5 9.7 9.9].

The parameter results are the trend of the mean time (Fig 18) or time standard deviation (Fig 19) it takes for ants to make decisions over the value of threshold standard deviation.

It can be see that, both mean time and time standard deviation, with the threshold standard deviation increasing, the trends are to first decline slightly, then stabilize, and finally rise significantly. This is because increasing the standard deviation appropriately can help the system quickly achieve its goals, while excessive standard deviation can reduce the stability of the system.

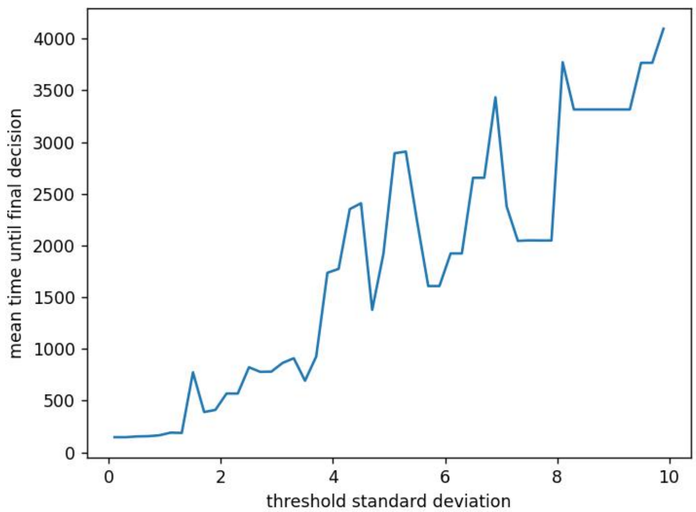

Figure 18. The trend of mean decision time over the value of threshold standard deviation

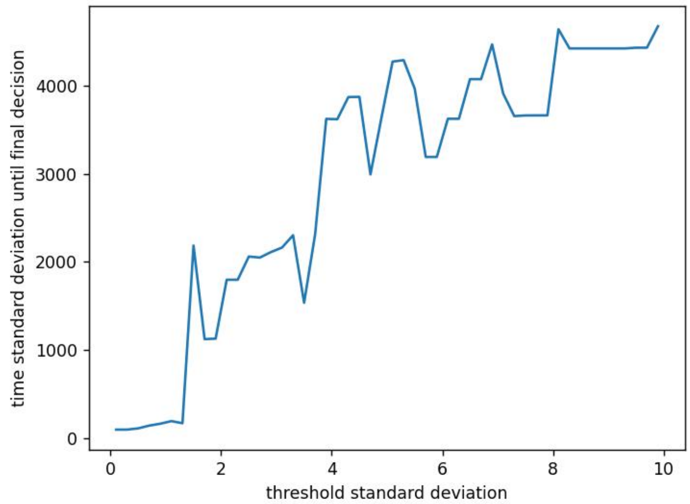

Figure 18. The trend of decision time standard deviation over the value of threshold standard deviation

References

[1] E. J. H. Robinson, N. R. Franks, S. Ellis, S. Okuda, and J. A. R. Marshall, 'A Simple Threshold Rule Is Sufficient to Explain Sophisticated Collective Decision-Making', *PLoS ONE*, vol. 6, no. 5, p. e19981, May 2011, doi: 10.1371/journal.pone.0019981.

[2] D. P. Kroese, T. Brereton, T. Taimre, and Z. I. Botev, 'Why the Monte Carlo method is so important today', *WIREs Computational Stats*, vol. 6, no. 6, pp. 386–392, Nov. 2014, doi: 10.1002/wics.1314.

Appendix

The altered code for read the pkl file and check the ant state

```
import pickle

with open('robinson_data1.pkl', 'rb') as file:
    state = pickle.load(file)

# current_time, discovers, visits, accepts, Ants, rnd_seed = state
print(len(state))
para_list = ['current_time','discovers','visits','accepts','Ants','rnd_seed']
```

```
for i in range(len(state)):
    print('------------------------------',para_list[i],'------------------------------')
    print(state[i])
```

The altered code in Q2

```
robinson.plot_summary_data_show_num_selected(current_time, accepts, discovers, visits, Ants)
```

The altered code in Q4

```
time_means = np.array([[1, 54, 54], [54, 1, 81], [54, 81, 1]])
probs = np.array([[0.70, 0.15, 0.15], [0.15, 0.70, 0.15], [0.15, 0.15, 0.70]])
```

The altered code in Q6

```
threshold_stddev_list = np.arange(0.1,10,0.2)
current_time_mean_list = []
current_time_std_list = []
print(threshold_stddev_list)

for threstd_i in threshold_stddev_list:
    state = robinson.run_model_simple(n, quals, probs, threshold_mean, threstd_i,
                                      qual_stddev, time_means, time_stddevs, [])
    current_time, discovers, visits, accepts, Ants, rnd_seed = state

    current_time_mean = sum(current_time)/len(current_time)
    current_time_mean_list.append(current_time_mean)

    diff_squared = [(x-current_time_mean) **2 for x in current_time]
    sum_diff_squared = sum(diff_squared)
    current_time_std = math.sqrt(sum_diff_squared/len(current_time))
    current_time_std_list.append(current_time_std)

print('current_time_mean_list', current_time_mean_list)
print('current_time_std_list', current_time_std_list)

plt.plot(threshold_stddev_list, current_time_mean_list)
plt.xlabel('threshold standard deviation')
plt.ylabel('mean time until final decision')
plt.show()
```

```
plt.plot(threshold_stddev_list, current_time_std_list)
plt.xlabel('threshold standard deviation')
plt.ylabel('time standard deviation until final decision')
plt.show()
```