Master Project on MCSS
Progress Report 1

Huy Phung
University of Konstanz

March 22, 2020

1 Introduction

We consider the collective action of a bee colony. Each bee in a colony could possibly sting after observing a threat in the surrounding environment, and warn other bees by releasing pheromone. By sensing the pheromone released in the environment, other bees in the colony may also sting. Since stinging leads to the termination of an individual bee, it reduces the total defense capability as well. We studies how the actions of a bee changes with regarding to its surrounding the environment.
In this progress report, the following points are being presented:

1. Description of the system
2. Probability distribution of the number of stinging bees
3. Markov chain formalization
4. Parameter inference.

2 Formal description

There are 3 assumptions on the system:

1. Each bee release an unit amount of pheromone immediately after stinging.

2. A bee dies after stinging and releasing pheromone. In the other words, no bee can sting more than once.
3. Stinging behaviour only depends on the concentration of pheromone in the environment.

Given N bees, let $s : \{1, \ldots, N\} \to [0, 1]$ be the probability of each bee sting given i units of pheromone in the environment.

The question we are concerning about is that, given a system, how many bees would sting in the steady state. Formally Y is a discrete random variable indicates the number of dead bees at steady state, we want to find the probability distribution of Y Given n bees in an isolated box. After stinging, each bee release unit Δ amount of pheromone and then dies. We denote that

- $s(i)$ is the probability of a bee sting at a pheromone level $i\Delta(i \in \mathbb{N})$
- At pheromone

Regarding the way we observe the deadbees, there are two types of experiment that will be described

- Synchronous
- Asynchronous

2.1 Fully asynchronous experiment

In asynchronous experiment we assume that there is almost improbable for two bees to sting at exactly the same time, and any bees release pheromone immediately after its death. By that observation we can assume that each bee sting at different level of pheromone.

Lemma 1 In *asynchronous model*, the probability of seeing j dead bees at the steady state is

$$P(Y = j) = \binom{n}{j} s(0)s(1)\ldots s(j-1)(1 - s(j))^{n-j} \tag{1}$$

Proof: Since under asynchrnous assumption, we assume that at each pheromone level there is at most one bee sting. Thus, j stinging bees must do so under

different concentration of pheromone. Also under asynchronous update, at steady state there are at most j amount of pheromone diffused in the environment. The other bees does not sting at any pheromone level from 0 to j, thus we can see all of them as do not stinging at the highest level j.
Selecting j bees from N, each of them sting under pheromone level 0 to j, then the other $n-j$ bees does not sting under pheromone j, we deliver Lemma 1 directly. □

2.2 Fully synchronous experiment

In *fully synchronous* experiment we assume that the number of stinging bees is only counted after a fixed amount of time, so that without loss of generality we can assume the pheromone diffuse almost immediately among the bee colony, and each bee decide to sting or not to sting immediately after sensing the pheromone concentration.
Under the synchronous assumption, the Lemma 1 does not hold anymore. It is due to the fact that it is possible to have more than one bee sting given the same concentration of pheromone in the environment. Thus, we would like to construct the probability distribution in a different way.
In order to construct a formula for fully synchronous setup, we use the following observations:

1. if there are k bees stinging, the amount of pheromone in the environment is k

2. there are k intervals of stinging corresponding to pheromone amount of k, namely $(s(0), s(1)), \ldots, (s(k-1), s(k))$

3. placing bees into intervals must be consecutive, namely there exist no empty interval $(s(i), s(j))$ such that $(s(i-2), s(i-1))$ and $(s(j+1), s(j+2))$ are non empty.

Based on these observations, we may propose to use the following counting scheme for k bees stinging: *How many ways to place k bees into consecutive intervals starting from $(s(0), s(1))$?*.
For a consecutive interval of size l, the number of ways to place k bees into the interval such that there is no empty interval is the Stirling number of the

Second type, namely:

$$\left\{ {k \atop l} \right\} = \frac{1}{l!} \sum_{i=0}^{l} (-1)^i \binom{l}{i} (l-i)^k$$

However, Stirling number of the Second kind assumes does not retain the information of which the partition (interval) labels, which we need to calculate the probability. We need to embed the information into our calculation, so that the probability of k bees stinging within l intervals is

$$\frac{1}{l!} \sum_{i=0}^{l} (-1)^i \binom{l}{i} (l-i)^k (1-s(i))^l$$

Huy: I am not sure about this step, namely how I add the information of the interval labels By that we calculate the probability that given n bees, k bees among them stinging:

$$P(Y=k) = \binom{n}{k} \sum_{i=0}^{k} \frac{1}{i!} \sum_{j=0}^{i} (-1)^j \binom{i}{j} (k-j)^i (1-s(i))^j$$

However, this construction does not leads to the correct answer, e.g $(1-s_0)^2$ for $P(Y=0)$.....

3 System modeling using Markov Chain

3.0.1 Discrete time Markov Chain

A *Discrete Time Markov Chain* is a tuple, where

- a
- b

Markov property

3.0.2 Modeling bees colony

4 Modeling with Markov chain

4.1 Parameterized DTMC

parameter inference problem is defined as [2]

4.2 Statistical parameter inference

4.2.1 Definition

Statistical method for parameter inference is presented in [1] Input: distribution on steady state of number of dead bees Output:

4.2.2 Method description

Observing

5 Bayesian inference

5.1 Bayesian parameter inference

Essentially, Bayesian formula is presented as follow

$$P(H|D) = \frac{P(D|H)P(H)}{P(D)}$$

Prior distribution
Posterior distribution
Marginal distribution

5.2 Posterior conjugation

5.2.1 Conjugation of binary distribution

Lemma:

5.2.2 Conjugation of multinomial distribution

5.2.3 Problem declaration

5.2.4 Method description

References

[1] Matej Hajnal et al. "Data-Informed Parameter Synthesis for Population Markov Chains". In: *International Workshop on Hybrid Systems Biology.* Springer. 2019, pp. 147–164.

[2] Joost-Pieter Katoen. "The probabilistic model checking landscape". In: *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science*. 2016, pp. 31–45.