UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING

MASTER THESIS no. 1368

Agent model based on hybrid stochastic automata

Patrik Kolarić

Zagreb, July 2016

UNIVERSITY OF ZAGREB
FACULTY OF ELECTRICAL ENGINEERING AND COMPUTING
MASTER THESIS COMMITTEE

Zagreb, 16 March 2016

MASTER THESIS ASSIGNMENT No. 1368

Student:	**Patrik Kolarić (0036468417)**
Study:	Electrical Engineering and Information Technology
Profile:	Control Engineering and Automation

Title: **Agent model based on hybrid stochastic automata**

Description:

Thesis considers development of agent model that moves in 2D space. The model should be based on stochastic hybrid automaton. The model parameters should be determined based on honeybee movement in closed 2D space. Influence of the model parameters (caused by changes in the environment) on the aggregation of multi-agent system should be explored. Validity of the model should be tested by simulation in Matlab and compared with honeybee behavior.

Issue date:	18 March 2016
Submission date:	1 July 2016

Mentor:

Full Professor Stjepan Bogdan, PhD

Committee Secretary:

Associate Professor Mario Vašak, PhD

Committee Chair:

Full Professor Stjepan Bogdan, PhD

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA
ODBOR ZA DIPLOMSKI RAD PROFILA

Zagreb, 16. ožujka 2016.

DIPLOMSKI ZADATAK br. 1368

Pristupnik: **Patrik Kolarić (0036468417)**
Studij: Elektrotehnika i informacijska tehnologija
Profil: Automatika

Zadatak: **Model agenta zasnovan na stohastičkom hibridnom automatu.**

Opis zadatka:

Radom na zadatku potrebno je izraditi model gibanja agenta u dvodimenzionalnom prostoru. Model treba biti zasnovan na stohastičkom hibridnom automatu. Parametre modela potrebno je identificirati iz hoda pčele u zatvorenom prostoru. Nadalje, potrebno je ispitati utjecaj parametara na agregaciju više agenata, uzrokovanu promjenama u okolini (zagrijavanje određenog dijela dvodimenzionalnog prostora). Valjanost modela potrebno je ispitati simulacijom u Matlabu uz različite vrijednosti identificiranih parametara i usporedbom sa snimkama hoda pčele.

Zadatak uručen pristupniku: 18. ožujka 2016.
Rok za predaju rada: 1. srpnja 2016.

Mentor:

Prof. dr. sc. Stjepan Bogdan

Djelovođa:

Izv. prof. dr. sc. Mario Vašak

Predsjednik odbora za diplomski rad profila:

Prof. dr. sc. Stjepan Bogdan

I would like to thank my mentor, prof. dr. sc. Stjepan Bogdan for support and motivation during my studies.

CONTENTS

1. Introduction

Behaviour of animal swarms is an example of a complex process with many excellent qualities that could enhance modern swarm robotics, artificial intelligence applications, telecommunications and many other engineering fields. Examples of intelligent swarms are ant colonies, bee colonies, bird flocking, animal herding, bacterial growth, fish schooling, etc. They all produce some form of optimal behaviour when engaging certain tasks of survival.
One of the main properties that are naturally assigned to biological swarms is the decentralized approach to problem solving. There is no central unit that controls and observes every aspect of the system and yet the system remains highly coherent through the complex self-organisation processes. The absence of a single centre of control makes system highly robust to the damages or failures of it's parts. However, if a central unit in a centralised system is not functioning properly due to some malfunction, the whole system is immediately non-functional. Server controlling pack of robots is an example of centralised system. Once that server is down, robots have no lead and they are useless until server is back on.
Secondly, the common property of swarms is that every agent as an individual has considerably simple tasks when compared to the task of the colony as a whole. The economical benefits of a system with such set-up are hard to overlook. System that comprises dozens of very cheap robots will in most cases be cheaper than the system composed of one central unit and few robot workers. It is the central unit what makes the system expensive, for it must guarantee extraordinary robustness to damages and maintain constant power supply and communication capabilities. If an agent in decentralised system is no longer functional, it would not perturb system significantly and the damaged agent could be easily replaced by a new one.

Endeavors of modern control engineers to demystify those complex behaviours in the form of algorithms are now even stronger as cutting edge technology allows better communication and faster realisations of the algorithms that once were too cumbersome

for a real-time applications. The main question is: How does the behaviour of a single agent affect the behaviour of the whole colony. Once that the answer to this question is acquired in the form of a mathematical algorithm, it is easy to mimic the colony behaviour by simply programming agents to follow the steps of the algorithm previously devised.
One such attempt is ASSISI project. ASSISI stands for Animal and robot Societies Self-organise and Integrate by Social Interaction ((T. Schmickl, 2013), (Griparic)). One of the main goals is to understand how animal swarms behave and to interact with this behaviour by using robot agents. In order to achieve that, agent dynamics must be modeled and it's influence on colony dynamics investigated. System observed is the closed arena containing bees and units that can produce various stimulus.

This thesis tackles the problem of simulating agent dynamics modeled by Beeclust Algorithm and predicting how parameters of single agent dynamics affect the overall dynamics. Arena with one heat center and bees as agents is simulated in MATLAB.
In the biological world, governed and driven by stochastic principles, it is nearly impossible to predict behaviour and to influence it in a deterministic domain. That is why probabilistic and stochastic mathematical tools are used in the thesis. Using Markov chains is a very common approach for modeling biological systems. For that reason it is used in this thesis to abstract agent's individual parameters to a macro-level characteristics of the whole colony.
Finally, mathematical predictions are compared with simulation to evaluate their accuracy.

2. Mathematical model of movement dynamics

The choice of mathematical model that describes dynamics of a bee agent should approximate behaviour of actual biological bee as close as possible. Even small inaccuracies in the equations of a single agent could produce considerable deviation of the multi-agent system behaviour from real colony behaviour. Therefore, every effort put into a closer observation of bees and obtaining more realistic mathematical model is rewarding.

Recently devised Beeclust algorithm (Hereford, 2011) is the usual choice when modeling behaviour of a single agent and is used in this thesis as well. The main advantage of the algorithm is its mathematical simplicity and compactness on the one hand and precise approximation of real bee behaviour on the other. Usage of more complex models could result in better simulations, however, attempts of devising macro level mathematical model from micro level models is usually quite involving. Therefore, simplicity of the mathematical model from Beeclust algorithm is quite convenient for purposes of further mathematical analysis which makes it appropriate for this thesis.

Upon close observation of bee colonies in the closed arena it became apparent that they can be classified into four distinct types. "Goal Finder" is a type that finds the optimal point in space(stimulus based optimum) in short time and stays there. "Wall Follower" stays close to the walls of the arena. "Immobile agent" stands still and barely moves. "Random Walker" moves randomly in the space with a slight gradient towards the optimal spot in the arena. "Random Walker" is the most common of four types and the hardest one to analyze. Randomness of its movements brings stochastic elements to the mathematical model.

Since other types are mostly deterministic and their numbers are minor compared to

"Random Walkers", it is sufficient to analyze only "Random Walkers" to get a broader picture of the colony behaviour. Thus, focus of this thesis is on global behaviour of the colony that consists of "Random Walkers", while other types are excluded from the analysis.

2.1. Beeclust Algorithm for random walkers

Movement of "Random Walkers" can be modeled by using Beeclust algorithm. What follows is pseudocode of the algorithm:

```
procedure BEECLUST
    STATE: MOVE (Initially, agent moves randomly in 2D space.)

    for time elapsed since start of simulation < simulation end time do
        if agent encounters another agent then

            Calculate waiting time

            STATE: WAIT
            for time from encounter < pre-caclulated waiting time do
                STATE: WAIT
            end for
        else
            STATE: MOVE
        end if
    end for
end procedure
```

Basically, what Beeclust algorithm says is that there are two possible agent's states: MOVE or WAIT. Both discrete states are characterised by corresponding system of differential equations.

2.1.1. State: MOVE

Agent will initially move around the arena in search for the optimal point. In case of bees, the optimal point will be warm spot, bright spot, spot with no mechanical vibrations (bees are very sensitive to the mechanical vibration and will tend to avoid it)... Since the choice of stimulus does not affect efficiency of the algorithm, only temperature stimulus is used in this thesis. Agent's tendency towards the optimal spot is modeled by using temperature gradient. Although there is a slight inclination towards temperature gradient, resulting direction of movement is mostly random.

Differential equation that describes 2D movement should account for both elements, randomness of direction and bias resulting from influence of temperature gradient. General form of equation for agent i is the following,

$$\frac{dx_i(t)}{dt} = \beta_i \cdot [\alpha_i \cdot (x_\tau - x_i(t)) + (1 - \alpha_i) \cdot \xi_i(\sigma, t)] \tag{2.1}$$

where parameter α_i describes the influence of temperature gradient relative to the influence of randomness in choice of direction.

$x_\tau - x_i(t)$ stands for temperature gradient, with x_τ being position of the heat source and $x_i(t)$ position of agent i at time t.

$\xi_i(\sigma, t)$ is a random variable ranging between $[-\pi, \pi]$. The correct choice of its probability distribution is determined by bee's natural tendency to change the direction of movement. It is intuitively clear that large, abrupt changes of direction are unusual for bees, as well as for the most living organisms. Bee will most probably move in the direction which its senses confirm to be safe. Senses provide much better information of the space in front of the bee than of the space behind it. For that reason, it is sensible to constrain the range of possible directions to a range smaller than $[-\pi, \pi]$. $[-3\pi/4, 3\pi/4]$ is presumed in this thesis. The most straightforward choice of probability distribution is the uniform distribution. However, in order for the model to resemble real distribution, a normal distribution should be used instead. Uniform distribution in range $[-3\pi/4, 3\pi/4]$ can be replaced by normal distribution with $\sigma = 3\pi/4$ resulting in more realistic random variable with similar probabilistic properties. In that case, around 70% (2σ) of values will fall within range $[-3\pi/4, 3\pi/4]$ with the most probable directions close to 0.

β_i is parameter that scales the norm of velocity according to agent's type. It describes overall mobility of an agent. In equation of "Immobile agent" dynamics, β_i takes small values. In case of other agents, β_i is close to 3 cm/s which is the mean velocity of 3 days old bees.

Previously mentioned set of equation that describes dynamics of bee's movement is general description and is somewhat different from the set realised in the simulation. Detailed description of equations used for simulations can be found in the section "Simulator".

2.1.2. State: WAIT

Whenever encounter occurs between two agents their state is changed to: WAIT. They stop and wait for a time determined by the presence of stimulus, temperature in this case. The closer temperature is to the optimal one, the greater is the tendency of bee to stay there for a longer time. If the spot is unpleasantly hot or cold, the bee will be more reluctant to stay. This behaviour of a single agent is essential property that enables aggregation of the whole colony at the optimal spot. Gradient alone is not sufficient for creation of stable clusters of bees.

The functional relationship between waiting time and temperature is found experimentally. One possible solution is:

$$T_{wait} = T_{wait,max} \cdot \frac{e^{\tau/6} - 1}{e^6 - 1} \tag{2.2}$$

Formally differential equation for dynamics in this state is:

$$\frac{dx_i(t)}{dt} = 0 \tag{2.3}$$

Time spent in WAIT state becomes a trigger for the next transition. At this moment, it is sensible to introduce a new variable for monitoring that time:

$$\frac{dw_i(t)}{dt} = -1 \tag{2.4}$$

$$w_i(0) = T_{wait} \tag{2.5}$$

where T_{wait} stands for waiting time. Equation 2.4 is a simple integrator and w_i a countdown clock variable.

As follows, complete behaviour of an agent in state WAIT can be compactly written as a system of two equations: 2.3 and 2.4 with initial condition 2.5

3. Mathematical background

Mathematical models of complex biological systems are usually characterized by considerably stochastic nature. This is especially the case for the micro-level systems. For instance, chemical reactions, genetics, enzyme kinetics from micro-level system and predator-prey process, species extinction, epidemic process, population genetics as examples from the macro world (Allen, 2010). All of them lacked analytical solutions until mathematical tools such as probability and stochastic were sufficiently developed. Stochastic theory proved to be significantly powerful for modeling not only biological systems, but economical, social and many others as well. Mathematical model of the system introduced in the last chapter indicates highly stochastic nature. It is for that reason sensible to suppose that the application of stochastic theory in analysis should be able to predict the behaviour of that system to a decent level of accuracy.

Discrete-time Markov Chains (DTMC) seemed to be the right choice for that attempt. There were requirements for the mathematical simplicity and computational applicability on the one side and relevance of the mathematical tool to the specific problem of bee colony movement prediction on the other. DTMC turned out to be the best trade-off option.

In addition to the short overview of DTMC, this chapter briefly covers important mathematical theorems used in solving the main problem of the thesis. Refer to following ((Allen, 2010) and (Elezović, 2007))for details on topic of the following chapter

3.1. Discrete-time Markov Chains

3.1.1. Markov Property

Important characteristic of the stochastic processes analysed in the following sections is that their future values do not depend of the past values. Therefore, given some present time, the value of a variable in the future step can be calculated without the knowledge of any values that variable held previously. It is also called a memorylessness.

Definition 1. ***Markov Property***

Stochastic process $X = x_1, x_2, ...$ has Markov Property if following holds:

$$Prob(X_n = x_n | X_{n-1} = x_{n-1}, ..., X_0 = x_0) = Prob(X_n = x_n | X_{n-1} = x_{n-1}) \quad (3.1)$$

The stochastic process that satisfies Markov Property is called Markov Process.

Definition 2. ***One-step transition probability***

One-step probability is the probability of transition from state i to state j in one discrete step:

$$p_{ji}(n) = Prob(X_{n+1} = j | X_n = i) \quad (3.2)$$

In general, every state i can be reachable from any state j with the probability p_{ij}. This can be compactly written in the form of the matrix called Transition Matrix:

Definition 3. ***Transition Matrix***

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1n} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & p_{n3} & \cdots & p_{nn} \end{bmatrix} \quad (3.3)$$

Every column j encompasses all possible transitions from state j (including "transition" to itself). Therefore, following holds:

$$\sum_{i=1}^{n} p_{ij} = 1 \tag{3.4}$$

In the case of a DTMC every transition can be one-step only. As follows, there are three possible transitions from state S_n in DTMC:

$$S_{n+1} = \begin{cases} S_n + 1, & p_{S_{n+1}S_n} \\ S_n - 1, & p_{S_{n-1}S_n} \\ S_n, & p_{S_nS_n} \end{cases} \tag{3.5}$$

Transition Matrix in the specific case of a DTMC is:

$$P = \begin{bmatrix} p_{11} & p_{12} & 0 & \dots & 0 \\ p_{21} & p_{22} & p_{23} & \dots & 0 \\ 0 & p_{32} & p_{33} & \dots & 0 \\ 0 & 0 & p_{43} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & p_{n-1n} \\ 0 & 0 & 0 & \dots & p_{nn} \end{bmatrix} \tag{3.6}$$

It can be noticed that the system's deterioration from state 1 to the lower state is not possible and similar conclusion can be made for the state n. These two states are specific and they represent boundaries of the system. Since the origin of the states in this thesis is geometrical (covered in next chapter), boundaries also represent geometrical restrictions of the system. In general, DTMC can have infinite number of states.

As its name indicates, Transition Matrix can be used to calculate transitions of system's states in a discrete time-step.

Definition 4. ***System transition***

If the vector p_n represents vector of states in a time instance n and P a Transition Matrix of that system then following holds:

$$p_{n+1} = P \cdot p_n \tag{3.7}$$

If p_0 is the initial state vector, then the state vector after an arbitrary number of time-steps N is:

$$p_N = P^N \cdot p_0 \tag{3.8}$$

Transition Matrix is a very informative mathematical object that can be easily used. Once it is obtained, analysis of system's probability distribution is an easy task of sequential matrix multiplication. In that case p_0 represents initial probability distribution of being in a certain state and p_N represents probability distribution of being in a certain state at time instance N. In a slightly different scenario, p_0 can also be deterministic number of some objects in corresponding state and p_N an estimation of numbers in each state for time instance N.

Transition Matrix is in the very core of the solution provided in this thesis. Following chapters deal with the construction of Transition Matrix by calculating corresponding probabilities.
Afterwards, it is used to estimate time necessary for the cluster creation and the degree of its integrity.

3.2. Other Mathematical Definitions

This section is a short overview of stochastic definitions used for construction of probabilities. Some fundamental laws and theorems of probability theory are not reviewed here.

3.2.1. Combinations

Definition 5. ***Combinations*** *The number of ways that k distinct objects can be picked out of n is:*

$$\binom{n}{k} \tag{3.9}$$

3.2.2. Algebra of Independent Random Variables

Random variable appears in the equations of the bee's movement dynamics and manipulating with these variables is not straightforward. The following definitions are covering all of the mathematical manipulations used for calculations of probabilities.

Definition 6. ***Sum of Continuous Random Variables***

Let X and Y be continuous random variables with density functions $f(x)$ and $g(y)$ respectively. Then the sum $Z = X + Y$ is a random variable with density function $f(z)$, where $f(z)$ is the convolution of $f(x)$ and $g(y)$:

$$(f * g)(z) = \int_{-\infty}^{\infty} f(z-y)g(y)dy \tag{3.10}$$

or:

$$(f * g)(z) = \int_{-\infty}^{\infty} g(z-x)f(x)dx \tag{3.11}$$

There is a good reason why normal distribution is the most frequently used probability density function in stochastic theory. Not only does it mimic nature's events quite efficiently, but is also very practical for mathematical manipulations. The following definition endorses this remark:

Definition 7. ***Sum of Continuous Normally Distributed Random Variables***

Let X and Y be continuous normally distributed random variables with corresponding parameters $\mathcal{N}(\mu_x, \sigma_x)$ and $\mathcal{N}(\mu_y, \sigma_y)$, where μ is expected value of Normal distribution and σ deviation of Normal distribution. Then the sum $Z = X + Y$ is also a normally distributed random variable with parameters: $\mathcal{N}(\mu_z, \sigma_z)$ where:

$$\mu_z = \mu_x + \mu_y \tag{3.12}$$

$$\sigma_z^2 = \sigma_x^2 + \sigma_y^2 \tag{3.13}$$

Instead of solving relatively complex convolution integral (3.10 and 3.11), the sum of normally distributed random variables can be easily found by using simple relations 3.12 and 3.13.

Definition 8. ***Function of Random Variable***

Let $Y = \psi(X)$. If ψ is increasing or decreasing function and $f(x)$ is a probability density function of random variable X then probability density function of random variable Y can be calculated as follows:

$$g(y) = f(x) \cdot \left|\frac{dx}{dy}\right| \tag{3.14}$$

$$y = \psi(x) \tag{3.15}$$

,

or alternatively:

$$g(y) = f(\psi(y)^{-1}) \left| \frac{d\psi(y)^{-1}}{dy} \right| \tag{3.16}$$

,

Definition 9. ***Expected Value***

Expected value of a continuous random variable X can be calculated using the following integral:

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx \tag{3.17}$$

,

This chapter sums up mathematical relations used in the next chapters for the probabilistic calculations and relations used for the application of these probabilities in the prediction of the colony behaviour.

4. Simulator

Instead of using general differential equation 4.4 of 2 dimensional velocity vector, simulator is simplified by considering only the change of direction, while amplitude is fixed to the predefined value of agent's average velocity:

$$v_{avg} = 3cm/s \tag{4.1}$$

Difference equation of a direction is then following:

$$\phi_i(t + \Delta T) = \phi_i(t) + \alpha_i \cdot \Delta_{\tau,i}(t) + (1 - \alpha_i) \cdot \phi_{\mathcal{N}}(t) \tag{4.2}$$

where $\phi_{\mathcal{N}}$ is random variable with normal probability distribution (expected value is 0, standard deviation is $3\pi/4$), $\Delta_{\tau,i}$ is difference between current direction of agent i and direction of gradient and ΔT is discrete time unit of the simulation. The new direction of each agent is calculated in every step of simulation. Resulting velocity vector at time t is then:

$$\frac{dx_i(t)}{dt} = v_{avg} \cdot e^{j \cdot \phi(t)} \tag{4.3}$$

and it is updated consequently with direction update.

Euler forward method is used for discretization of equations in the simulator. The choice of discretization method is not a big issue when simulating slow macro biological systems like colonies of bees. Discrete time unit that is sufficiently small to correct for errors of discretization is easily achievable. Euler forward discretization method gives the following for position of an agent in the next time step:

$$x_i(t + \Delta T) = x_i(t) + \Delta T \cdot v_{avg} \cdot e^{j \cdot \phi(t)} \tag{4.4}$$

When an agent hits the wall of the arena its direction is automatically set perpendicular to that wall.

Size of the arena, discrete unit of time, number of agents, α, temperature distribution, waiting time function are all parameters that are set before the start of simulation. Guidelines for choosing parameters in accordance with true honeybee behaviour and realistic temperature distributions are found in ((Michael Bodi, 2012), (Thomas Schmickl, 2009) and (M. Szopek*))

Gaussian function is used for modeling temperature distribution. Solution of heat equation 4.5 describes the time-space propagation of the temperature distribution.

$$\frac{\partial \tau}{\partial t} = h^2 \left(\frac{\partial^2 \tau}{\partial x^2} + \frac{\partial^2 \tau}{\partial y^2} + \frac{\partial^2 \tau}{\partial z^2} \right) \tag{4.5}$$

In the steady state, a single source of heat will produce a temperature space distribution in the form of a Gaussian function:

$$\tau_i = \tau + (\tau_{casu} - \tau) \cdot e^{-(\frac{(\Delta x)^2}{2{\sigma_x}^2} + \frac{(\Delta y)^2}{2{\sigma_y}^2})} \tag{4.6}$$

where τ is a temperature of the space surrounding arena and τ_{casu} is the temperature of a source. σ_x and σ_y are not equal in general, but in this thesis, temperature distribution is uniform along every concentric circle which has center in the source of heat and thus, they are equal.

First spatial derivative of the temperature distribution gives a temperature gradient which has an effect on the movement of an agent:

$$\nabla_x = (\tau_{casu} - \tau) \cdot e^{-(\frac{(\Delta x)^2}{2{\sigma_x}^2} + \frac{(\Delta y)^2}{2{\sigma_y}^2})} \cdot \frac{\Delta x}{{\sigma_x}^2} \tag{4.7}$$

$$\nabla_y = (\tau_{casu} - \tau) \cdot e^{-(\frac{(\Delta x)^2}{2{\sigma_x}^2} + \frac{(\Delta y)^2}{2{\sigma_y}^2})} \cdot \frac{\Delta y}{{\sigma_y}^2} \tag{4.8}$$

Now, $\Delta_{\tau,i}(t)$ from 4.2 can be written as,

$$\Delta_{\tau,i}(t) = \phi_{\tau,i}(t) - \phi_i(t) \tag{4.9}$$

with $\phi_{\tau,i}(t)$ being,

$$\phi_{\tau,i}(t) = atan2(\Delta_y / \Delta_x) \tag{4.10}$$

5. Probability distribution analysis

This chapter elaborates the calculation of probabilities and creation of the transition matrix (3.6) for a specific case of a single agent in the arena at a certain distance from the heat source. It shows how parameters of simulation known beforehand and geometrical properties of the arena are used in calculation of the probabilities.

The heat distribution, size of the arena and the number of agents are global parameters that affect the probability distribution of agents during simulation. Individual parameters of an agent such as sensitivity to heat and corresponding waiting time, the degree of randomness in agent's movement modeled by alpha and velocity are also to be considered in the calculation.

5.1. Design of Markov States

The main idea of analysis is to interpret 2D space in such a way that Markov chain theory can be used in the analysis. The idea is similar to one introduced in (Hereford, 2011) where every cell in 2D space represents a single Markov chain.
Here, space is split into concentric disks around the center of heat as shown on Figure 5.1.

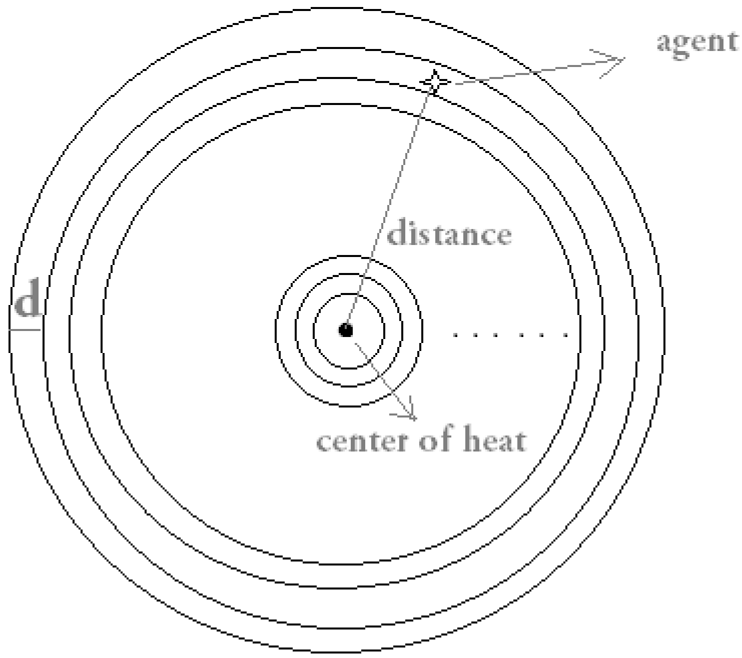


Figure 5.1: Concentric disks in the arena

Analysis becomes symmetrical and easier for implementation. Every disk is at certain distance from the center. Since Gaussian heat distribution is symmetrical around the center, it follows that the temperature inside the disk is approximately constant. Waiting time, which is a function of temperature, is constant as well. Temperature gradient of a bee inside the disk is always perpendicular to the smaller of two circles that surround the disk.

Now that it is clear that every disk has unique properties that are constant inside the disk, the question is how to incorporate that in the Markov chain theory.
One way of doing that is to define a single agent as a Markov chain or Markov process. Disk number represents the current state of an agent. When an agent is in zero state, it can only stay in the zero state or transition to the state 1. Zero state represents the center of the arena. State N represents the most distant disk in the arena.

When an agent is in the state i it has certain propensity to enter state $i + 1$, to enter state $i - 1$ or to stay in the state i This propensities are modeled by transition

probabilities which eventually form the transition matrix.
Probabilities of transition can be calculated based on systems of equations that model agent's movement, equations 4.4 and 4.2 when agent is in MOVE state and equation 2.3 when in WAIT state.

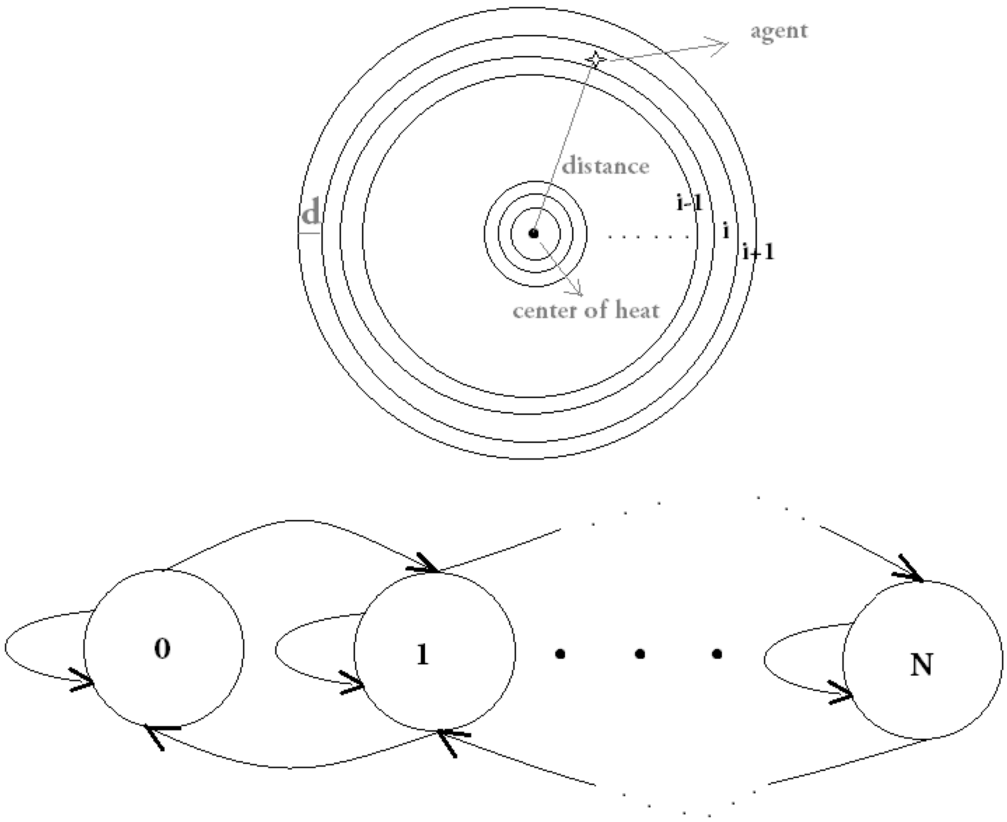


Figure 5.2: Concentric disks in the arena

5.2. Transition Probabilities

Input vectors for calculation of direction update (eq. 4.2) are shown on Figure 5.3. $\phi_i(t)$ is angle of agent's current direction, $\Delta_{\tau,i}$ is angle of rotation that matches vector of current direction with temperature gradient. These two angles are deterministic.

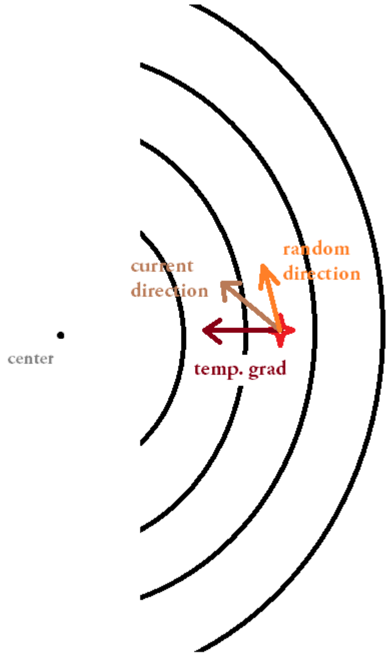


Figure 5.3: Vectors used in calculation of direction update

Third angle in calculation is random variable with normal distribution:

$$\phi_{\mathcal{N}}(t) \sim \mathcal{N}(\mu_\phi, \sigma_\phi) \tag{5.1}$$

$$\mu_\phi = 0 \tag{5.2}$$

σ_ϕ from 5.1 is tested in the last chapter for different values that are ranging between $2\sigma_\phi = [\pi, 2\pi]$, which assures that 95% of values falls in that range.

It can be noticed that curvature of larger circles that surround an agent is not as pronounced as curvature of circles closer to the center. In fact, for the most circles this curvature can be omitted which makes circles easily replaceable by straight lines (5.4). For a sufficiently large arena, this approximation holds from the agent's perspective. The group of circles that are not subject to this approximation are all considered cluster area anyway. The colony dynamics inside this small area of cluster is not as interesting nor significant as the event of aggregation itself. Since aggregation develops from the outside towards the center, this approximation should be satisfactory.

Velocity of every agent in simulation is constant. It only changes direction of movement in every discrete step. It means that agent will update its position on any of the points that lie on the circle around current position, with $r = \Delta T \cdot v_{avg}$ (Figure 5.4). The question is: What is the probability of moving to a certain point or to a certain state?

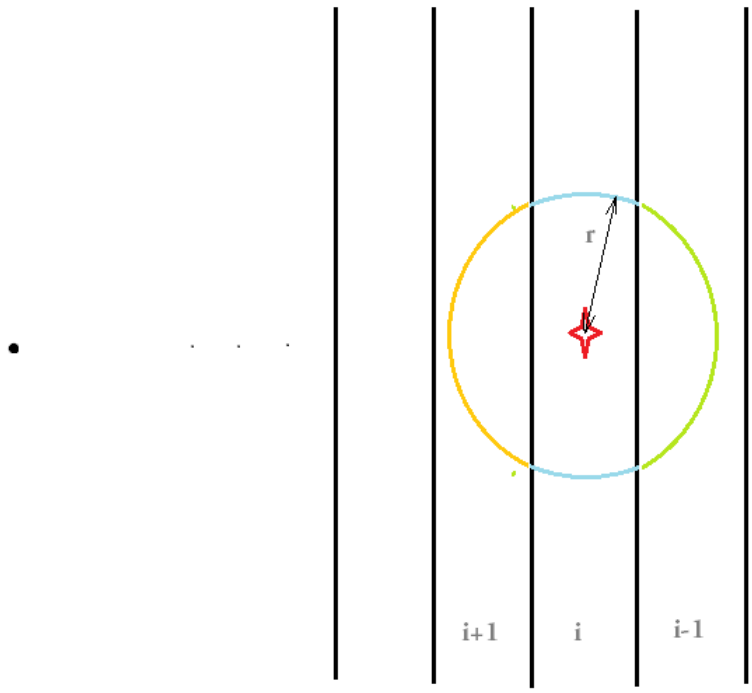


Figure 5.4: States approximated with horizontal disks

5.2.1. Width of Disks

Reach of agent's movement in one discrete time unit is defined by the length of that time unit and the agent's velocity. However, based on the width of disks agent may reach one of the neighbouring disks or may not leave current disk, all three events are happening with some probability. If width of disk is not chosen properly, agent may not be able to enter neighbouring disks or in the opposite scenario may be able to cross over the whole neighbouring disk, reaching farther disks which would be against Markov property assumption. Fortunately, the width of disks is not coupled with any parameter or variable in simulation and can be chosen in the way suitable for further analysis.

One of the assumptions is that the expected position of an agent inside a disk is on the center line of the disk (Figure 5.4). In order for the analysis to comply with that assumption, position update in the event of transition must on average assure that agent's new position is on the center line of the neighbouring disk.
For the sake of simplicity, probability distribution of direction update is assumed to be uniform $\phi \sim \mathcal{R}(-\pi, \pi)$. It is clear from Figure 5.5 that the probability distribution of direction update in the event of transition $(i+1)->i$ is $\phi \sim \mathcal{R}(-\theta, \theta)$.

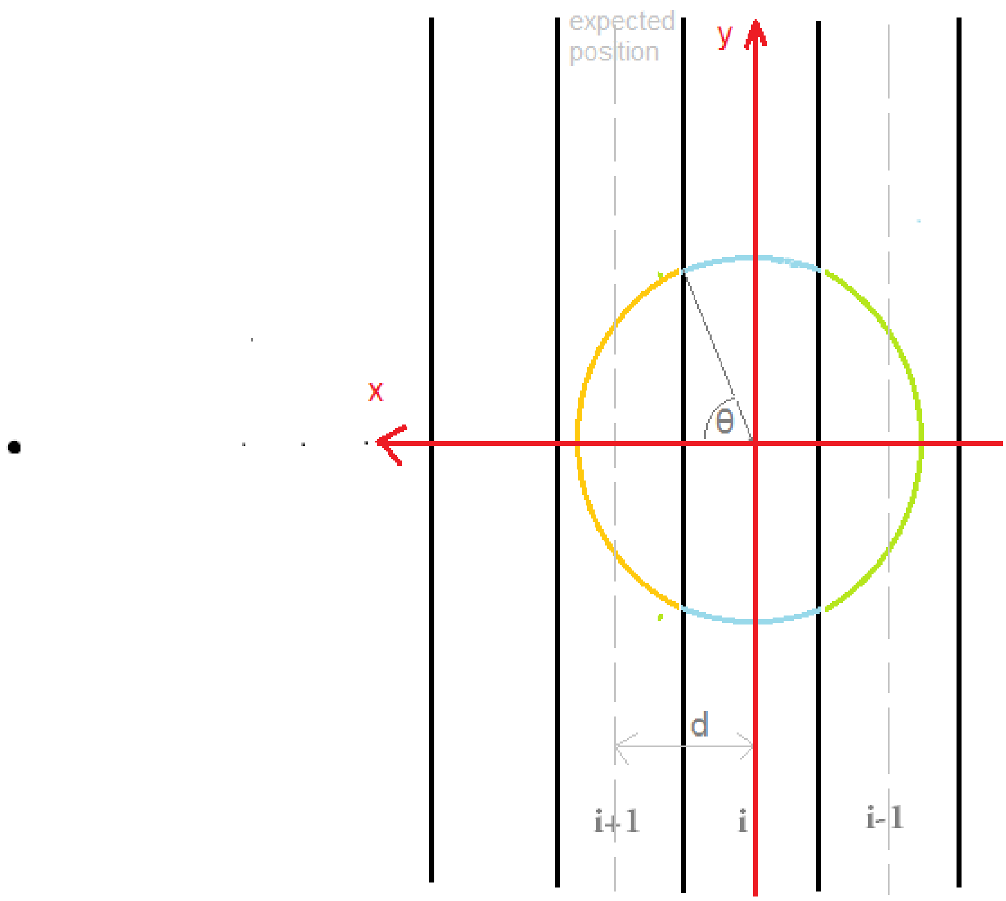


Figure 5.5: Calculation of θ and d

The question is how to determine the value of d. One way is to find θ such that expected value of x is d. In general, x is random variable X and its expected value can be found by using 3.17. Probability density function $f(x|i+1->i)$ must be determined first. Since x is a function of ϕ (Figure 5.5):

$$x = r \cdot cos(\phi) \tag{5.3}$$

then

$$\phi = acos(\frac{x}{r}) \tag{5.4}$$

Now that functional relation between x and ϕ is known it is possible to find probability density function of random variable Φ by using the Equations 3.14, 3.15 and 3.16.

$$f(x|i->i+1) = \frac{1}{\theta \cdot \sqrt{(1-x)^2}} \tag{5.5}$$

Expected value of $f(x|i->i+1)$ is:

$$\begin{aligned} E[X|i->i+1] = \int_{-\infty}^{\infty} & xf(x|i->i+1)dx = \\ \int_{\frac{d}{2}}^{r} & xf(x|i->i+1)dx = \\ & \frac{1}{\theta} \cdot [\sqrt{1-\frac{d^2}{4}} - \sqrt{1-r^2}] \end{aligned} \tag{5.6}$$

In order for probability analysis to meet the Markov property following must hold:

$$E[X|i->i+1] = \frac{1}{\theta} \cdot [\sqrt{1-\frac{d^2}{4}} - \sqrt{1-r^2}] = d \tag{5.7}$$

From Figure 5.5 follows:

$$\frac{d}{2} = r \cdot cos(\theta) \tag{5.8}$$

Solution to system of equations 5.7 and 5.8 gives the desired values of d and θ:

$$\theta = 1.1655 \tag{5.9}$$

$$d = v_{agent} \cdot \Delta T \cdot cos(1.1655) \tag{5.10}$$

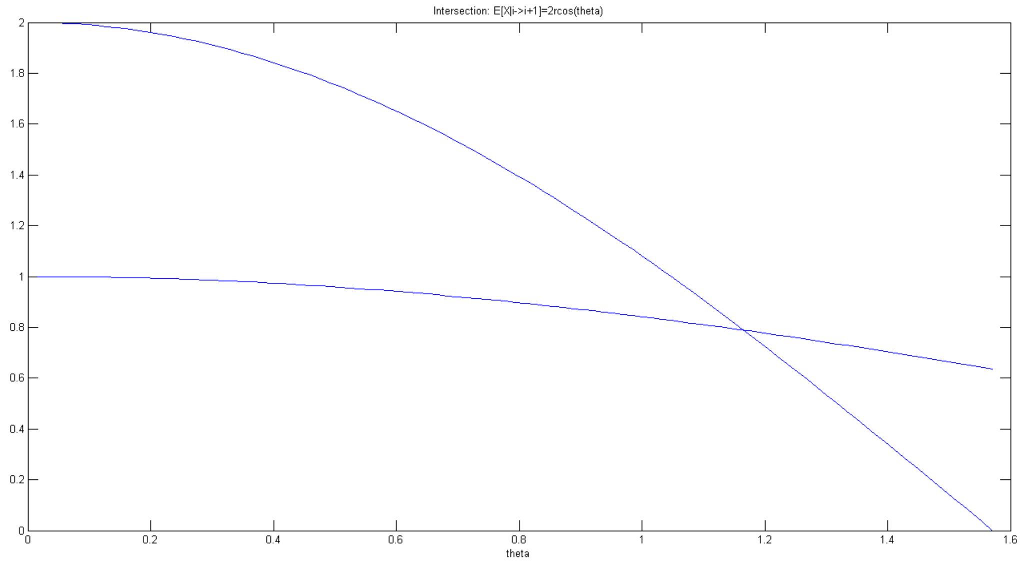


Figure 5.6: Solution for theta

θ has the same value in every simulation, while d is determined according to the values of v_{agent}, ΔT and θ. It is important to notice how three parameters affect the value of d. Smaller discretization time will produce thinner disks, greater in number compared to larger discretization time, which naturally makes analysis become more detailed. Still, overly detailed analysis is not necessary in the case of slow process such as bees movement.

5.2.2. Distribution of $\phi(k)$

Once that ϕ is statistically described in the form of probability distribution function, calculation of transition probabilities becomes trivial problem since it depends solely on direction of an agent. In this subsection approximation of ϕ is derived using difference equation algebra and statistical properties of the variables in the equation of direction update (Eq. 4.2).

Direction of an agent at some arbitrary initial moment is introduced as $\phi(0)$. In stochastic terms $\phi(0)$ is absolutely unknown and unbiased variable and can take any value in the range $[-\pi, \pi]$ with equal probability. It is convenient to describe it as:

$$\phi(0) = \mathcal{U}(-\pi, \pi) \tag{5.11}$$

From 4.2 and 4.9 it follows that $\phi(1)$ is:

$$\phi(1) = \phi(0) + \alpha \cdot \Delta_\tau(0) + (1-\alpha) \cdot \phi_\mathcal{N}(0)$$

$$\phi(1) = \phi(0) + \alpha \cdot (\phi_\tau(0) - \phi(0)) + (1-\alpha) \cdot \phi_\mathcal{N}(0) \tag{5.12}$$

$$\phi(1) = (1-\alpha) \cdot \phi(0) + \alpha \cdot \phi_\tau(0) + (1-\alpha) \cdot \phi_\mathcal{N}(0)$$

$\phi(2), \phi(3), ...$ can be determined in the same way. Finally, $\phi(k+1)$ has the following form:

$$\begin{aligned} \phi(k+1) = (1-\alpha)^{k+1} \cdot \phi(0) + \\ +\alpha \cdot [(1-\alpha)^k \cdot \phi_\tau(0) + (1-\alpha)^{k-1} \cdot \phi_\tau(1) + ... \\ ... + (1-\alpha) \cdot \phi_\tau(k-1) + \phi_\tau(k)] + \\ +[(1-\alpha) \cdot \phi_\mathcal{N}(k) + (1-\alpha)^2 \cdot \phi_\mathcal{N}(k-1) + ... \\ ... + (1-\alpha)^k \cdot \phi_\mathcal{N}(1) + (1-\alpha)^{k+1} \cdot \phi_\mathcal{N}(0)] \end{aligned} \tag{5.13}$$

$\phi(k+1)$ from the expression above depends on three functionally distinct terms.
$(1-\alpha)^{k+1} \cdot \phi(0)$ presents the effect that direction distribution at initial moment 0 has on direction distribution at time $k+1$ discrete time units after the initial moment. If the time lapsed from initial moment to moment k is large enough, this term can be neglected ($0 < 1-\alpha < 1$):

$$lim_{k->\infty}(1-\alpha)^{k+1} \cdot \phi(0) = 0 \tag{5.14}$$

$\alpha \cdot [(1-\alpha)^k \cdot \phi_\tau(0) + (1-\alpha)^{k-1} \cdot \phi_\tau(1) + ... + (1-\alpha) \cdot \phi_\tau(k-1) + \phi_\tau(k)]$ describes how each of the last k temperature gradients affects direction at moment

$k+1$. Older temperature gradients have smaller coefficients, which can be interpreted as memory loss. There is some time span, $N_T \cdot \Delta T$ that encompasses all the values of temperature gradients that have considerable effect on the value of current direction update $\phi(k+1)$. The temperature gradient depends only on the position of an agent since the heat center is assumed to be on the same position all the time. The whole term can be approximated by the random variable based on the area that is reachable from the position of an agent at time $k \cdot \Delta T$. All N_T positions that are taken into consideration must be inside the circle of radius $R = N_T \cdot v_{agent} \cdot \Delta T$.

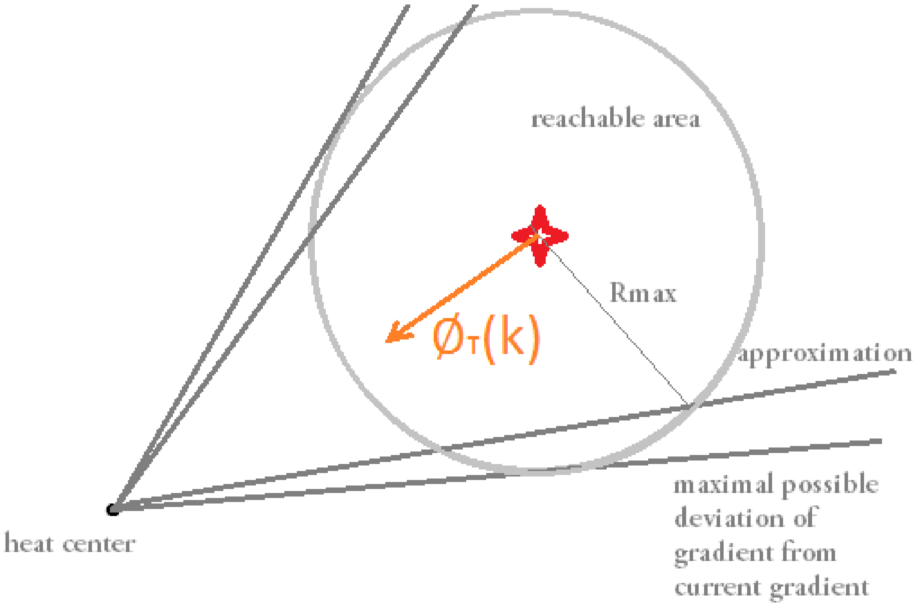


Figure 5.7: Approximation of gradient distribution

Crude approximation is to model the position distribution as a normal distribution, with points that are on the circle having very low probability of occurrence. Additionally, points on the edge of the circle are weighted with the smallest coefficients because they occur much before the central position.

Direction probability distribution can as well be approximated with Normal distribution. Simple way is to define most deviant gradients on the edge as $3 \cdot \sigma$ values of Normal distribution and current, central gradient as expected value. The maximum deviation is:

$$\phi_{\tau,DEV} = atan(\frac{v_{agent} \cdot \Delta T \cdot N_T}{l}) = atan(\frac{R}{l}) \tag{5.15}$$

where l is distance between the agent and the heat center. Now, whole term can be approximated with random variable Φ_τ :

$$\Phi_\tau \sim \mathcal{N}(\mu = \phi_\tau(k), \sigma = \frac{atan(\frac{v_{agent} \cdot \Delta T \cdot N_T}{l})}{3}) \tag{5.16}$$

Lastly, $[(1-\alpha) \cdot \phi_{\mathcal{N}}(k) + (1-\alpha)^2 \cdot \phi_{\mathcal{N}}(k-1) + (1-\alpha)^k \cdot \phi_{\mathcal{N}}(1) + (1-\alpha)^{k+1} \cdot \phi_{\mathcal{N}}(0)]$ represents randomness of the direction $\phi(k)$ as a result of randomness of all its previous instances. Again, not all previous instances have to be taken into consideration but only last N_T significant ones. The term is basically sum of random variables with Normal distribution weighted with powers $(1-\alpha)$ and can be calculated by using 3.12 and 3.13. Random variable with Normal distribution multiplied by positive coefficient smaller than 1 simply shrinks the distribution ($\sigma -> (1-\alpha)^i \cdot \sigma$). σ of distribution can be found by using 5.1 as:

$$\sigma^2 = ((1-\alpha) \cdot \sigma_{\mathcal{N}})^2 + ((1-\alpha)^2 \cdot \sigma_{\mathcal{N}})^2 + ... + ((1-\alpha)^{k+1} \cdot \sigma_{\mathcal{N}})^2 \tag{5.17}$$

The approximation of the whole term in the form of random variable is then:

$$\Phi_{\mathcal{N}} \sim \mathcal{N}(\mu = 0, \sigma = \sqrt{\sum_{n=1}^{k+1} \sigma_{\mathcal{N}}^2 \cdot (1-\alpha)^{2n}}) \tag{5.18}$$

Finally, the whole expression 5.13 is approximated as:

$$\phi(k+1) = \Phi_\tau + \Phi_{\mathcal{N}} \tag{5.19}$$

or

$$\Phi(k+1) \sim \mathcal{N}(\mu = \phi(k), \sigma = \sqrt{(\frac{1}{3} \cdot atan(\frac{R}{l}))^2 + \sum_{n=1}^{k+1} \sigma_{\mathcal{N}}^2 \cdot (1-\alpha)^{2n})}) \tag{5.20}$$

Taking last 10 seconds of history into consideration when calculating direction update probability distribution function, while assuming that $\Delta T = 0.05s$ or less, is sufficiently precise. The area that agent covers in that time is not too large which makes 5.15 a good approximation (agents in general do not change their position drastically in that time and thus hold relatively constant value for gradient).

Now, as direction update is known in the form of random variable with corresponding probability distribution function, the task of finding probabilities of a transition is an easy one.

5.2.3. Probabilities Calculation

If probability distribution of direction update is known, probabilities can then be calculated by finding probabilities of random variable falling in the range that corresponds to a certain state. Refer to the Figure 5.5. It is clear that direction of temperature gradient is collinear with x-axis. The probabilities of transition are following:

$$\begin{cases} i->i+1, & p_{i->i+1} = P(-\theta < \phi(k+1) < \theta) \\ i->i-1, & p_{i->i-1} = P(\pi - \theta < \phi(k+1) < \pi \\ i->i, & p_{i->i} = 1 - p_{i->i+1} - p_{i->i-1} \end{cases} \tag{5.21}$$

$$\begin{cases} p_{i->i+1}, & P(-\theta < \phi(k+1) < \theta) = \int\limits_{-\theta}^{\theta} f(\phi) d\phi \\ p_{i->i-1}, & P(\pi - \theta < |\phi(k+1)| < \pi) = \int\limits_{-\pi}^{-\pi+\theta} f(\phi) d\phi + \int\limits_{\pi-\theta}^{\pi} f(\phi) d\phi \\ p_{i->i}, & p_{i->i} = 1 - p_{i->i+1} - p_{i->i-1} \end{cases} \tag{5.22}$$

Where $f(\phi)$ is probability distribution function of direction update random variable (5.20).

Integral limits in expression 5.22 are written in the way that indicates geometrical representation of state transitions. It must be noted that $f(\theta)$ takes values beyond

$[-\pi, \pi]$ range, but all those values are aliases and are mapped to the original $[-\pi, \pi]$ range.

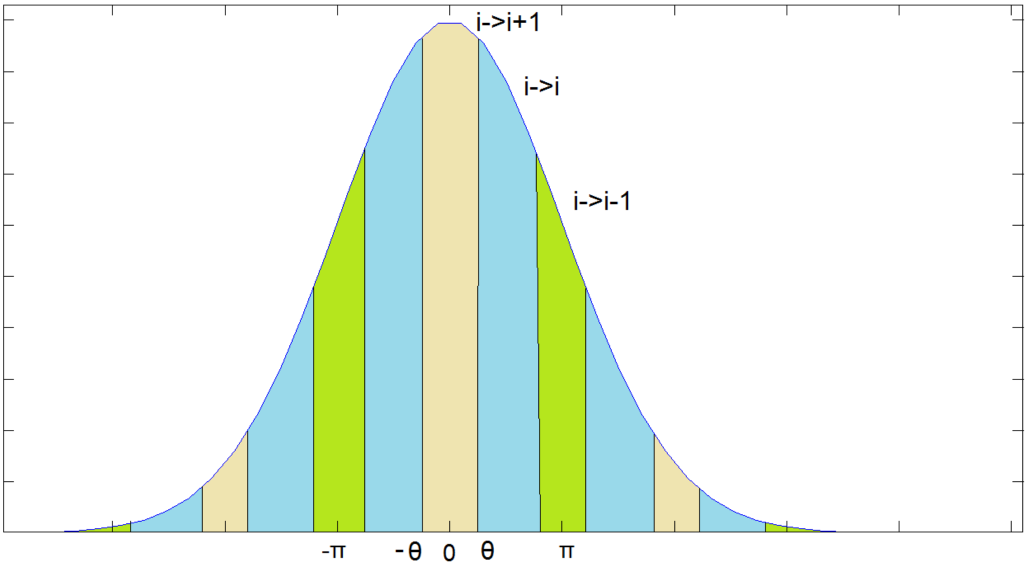


Figure 5.8: Gauss aliases, visual description

Finally, transition matrix is found by using 5.22 :

$$P = \begin{bmatrix} p_{1->1} & p_{1->2} & 0 & \dots & 0 \\ p_{2->1} & p_{2->2} & p_{2->3} & \dots & 0 \\ 0 & p_{3->2} & p_{3->3} & \dots & 0 \\ 0 & 0 & p_{4->3} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & p_{n-1->n} \\ 0 & 0 & 0 & \dots & p_{n->n} \end{bmatrix} \tag{5.23}$$

5.3. Waiting Time

Temperature gradient cannot keep the integrity of colony, but serves only to set a direction towards a potential cluster. It is waiting time, a parameter that models agent's inertia, what makes clusters possible. Optimal points in space should make agent highly inert, especially if that spot is picked through cooperative agreement of some considerable number of agents.
Mathematical description of waiting time used in simulation was introduced in previous chapters. This section attempts to abstract waiting time in the form of probability, similarly to what was done in previous section. Ultimately, the waiting probability is combined with previously derived probabilities of state transition.

If agent is positioned in an arbitrary point in the space, what is the probability of agent not moving? One approach is to use basic probability theory:

$$p[state = wait] = p[agent\ inside\ prZone] \cdot p[state \ = \ wait|agent\ inside\ prZone] \tag{5.24}$$

5.3.1. Exponential Distribution of Waiting

This subsection deals with probability of waiting under assumption that at least one agent is already inside the private zone, factor on the right side in 5.24.
If observed agent is positioned in the state (disk) i, then waiting time inside this disk can be approximated by taking value that corresponds to the central line, $T_{wait,i}$. What is the probability that agent stands still in current discrete moment? Refer to the Figure 5.9

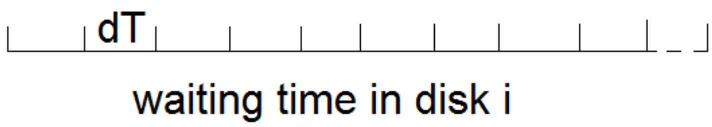


Figure 5.9: Calculation of waiting probability

Figure shows discretization of waiting time. First discrete ΔT unit after waiting time from Figure 5.9 corresponds to state MOVE. Out of $\frac{T_{wait,i}}{\Delta T} + 1$ time units, one corresponds to MOVE and others to WAIT. Rate of transition to state MOVE is then $\frac{1}{T_{wait,i}}$. This rate is valid only when agent is in state WAIT, namely has neighbours inside private zone. Once that rate of an event is known, it's probability can be found as:

$$p_i(next\ state = MOVE|current = WAIT) = 1 - e^{rate \cdot \Delta T} = \\ 1 - e^{-\frac{\Delta T}{T_{wait,i}}} \tag{5.25}$$

Similarly:

$$p_i(next\ state = WAIT|current = WAIT) = \\ 1 - p_i(next\ state = MOVE|current = WAIT) = e^{-\frac{\Delta T}{T_{wait,i}}} \tag{5.26}$$

5.3.2. Neighbour Existence Probability

More complex for calculation is the factor on the left (5.24). The reason for it's complexity is that it is not static, but changes in time (more precisely, it changes with probability distribution of agents in the arena, which is in functional relation with time). It is intuitively clear that the probability of existence of neighbours depends on current distribution of agents in neighbouring disks.

Luckily, the spatial probability distribution of one agent over the whole arena at the current moment is known and can be used to calculate the probability of having neighbour in private zone. Firstly, potential disks candidates for hosting neighbours are determined, based on the radius of private zone. To simplify calculation, all disks that are enclosed by private zone are treated equally, meaning that all potential disks contribute with the same percentage of their probability. This approximation is quite practical, since the observed agent can be anywhere around it's disk and neighbours can be anywhere around their hosting disks. Such property of symmetry allows to calculate probability of ONE agent being in private zone, p_{aIn} as sum of probabilities of corresponding potential neighbouring disks.

However, there are $N-1$ potential neighbours in the arena. Taking that into consideration gives following for the probability of one existing neighbour:

$$p_{ne,1} = \binom{N-1}{1} \cdot p_{aIn} \cdot (1 - p_{aIn})^{N-1} \tag{5.27}$$

In general:

$$p_{ne,k} = \binom{N-1}{k} \cdot p_{aIn}^{k} \cdot (1 - p_{aIn})^{N-k} \tag{5.28}$$

As follows,

$$\begin{gathered} p_{neExist} = \sum_{k=1}^{N-1} p_{ne,k} = \sum_{k=0}^{N-1} p_{ne,k} - p_{ne,0} = \\ (p_{ne,k} + (1 - p_{ne,k}))^{N-1} - p_{ne,0} = 1 - (1 - p_{aIn})^{N-1} \end{gathered} \tag{5.29}$$

Finally, probability of waiting can be derived by using 5.24, 5.26 and 5.29:

$$p[state = wait] = e^{-\frac{\Delta T}{T_{wait,i}}} \cdot (1 - (1 - p_{aIn})^{N-1}) \tag{5.30}$$

and

$$p[state = move] = 1 - p[state = wait] \tag{5.31}$$

5.3.3. Transition Probabilities and Waiting Time

Waiting time affects transition probabilities. Disks with small waiting time keep approximately same probabilities as found in Section 5.2. Disks with large waiting time however, become more absorbing by reducing outflow probabilities and increasing self-loop transition probability.
Formally, probabilities are found by combining 5.31 and 5.22:

$$\begin{gathered} \begin{cases} p_{i->i+1}, & P(-\theta < \phi(k+1) < \theta) \cdot P(state = move) \\ p_{i->i-1}, & P(\pi - \theta < |\phi(k+1)| < \pi) \cdot P(state = move) \\ p_{i->i}, & p_{i->i} = 1 - p_{i->i+1} - p_{i->i-1} \end{cases} \\ \begin{cases} p_{i->i+1}, & \int\limits_{-\theta}^{\theta} f(\phi) d\phi \cdot (1 - e^{-\frac{\Delta T}{T_{wait,i}}} \cdot (1 - (1 - p_{aIn})^{N-1})) \\ p_{i->i-1}, & (\int\limits_{-\pi}^{-\pi+\theta} f(\phi) d\phi + \int\limits_{\pi-\theta}^{\pi} f(\phi) d\phi) \cdot (1 - e^{-\frac{\Delta T}{T_{wait,i}}} \cdot (1 - (1 - p_{aIn})^{N-1})) \\ p_{i->i}, & p_{i->i} = 1 - p_{i->i+1} - p_{i->i-1} \end{cases} \end{gathered} \tag{5.32}$$

6. Testing

This chapter is the verification of analysis presented in the previous chapter. Different parameters are altered in a certain range of values and impacts of this alterations on the aggregation speed and intensity are observed.
System simulation and probabilistic simulation governed by Transition matrix are run parallel in time.
System simulation counts the number of agents in each disk at every discrete moment. Since there are much more disks than agents in the simulation, sampled data is extremely spiky (most disks do not have any agent at all, which exaggerates those disks that do have agents inside). Low pass filter is used to smooth the data and to plot descriptive probability distribution (or more precisely, distribution of number of agents).
Probabilistic simulation is absolutely independent of system simulation, except that they are synchronized in time. Probability distribution is found as:

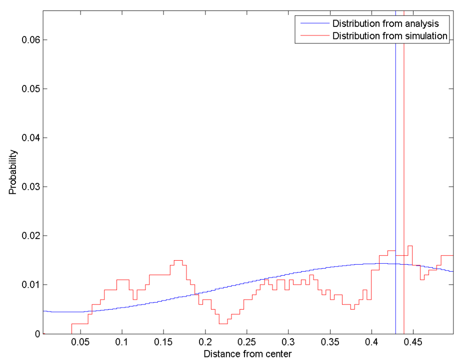
$$p(k) = P^k \cdot p(0) \tag{6.1}$$

where P is Transition matrix (3.6).
$p(k)$ from system simulation and from probabilistic simulation are shown together on plot. Number of disks is:

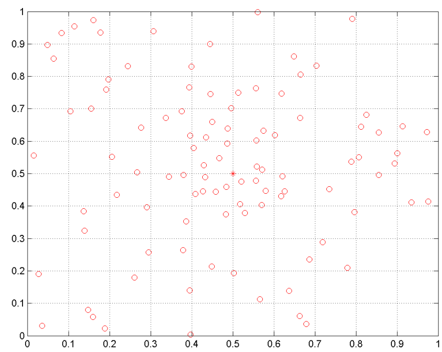
$$N_{disks} = \frac{R_{arena}}{d} \tag{6.2}$$

The power of analysis developed in previous chapter is that it gives probability distribution at any time during simulation. Whether the event of clustering happens or not is a matter of how the cluster is defined. Simple and effective way to describe a cluster is to use two parameters. One parameter is the minimal percentage of agents that can form the cluster. The other is the radius of disk that encloses the area of potential cluster.

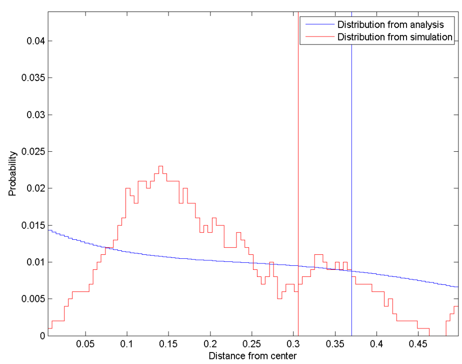
Set of Figures 6.1 shows the propagation of probability distribution in time for the following parameters: $\alpha = 0.085$, $\Delta T = 0.05s$, $random\ angle = [-pi/2, pi/2]$, $v_{agent} = 2.5cm/s$. Waiting time is not considered in this simulation, meaning that it's value is zero. Section 6.3 explores results of simulations with waiting time considered. Figures from the left column show agent's probability distributions, where vertical lines indicate disks that enclose 80% of agents. Right column shows the actual positions from simulation. Even though the experiments with agents in arena may not comprise more than 20-30 agents, simulation in this thesis runs with 100 agents. Probabilistic properties and conclusions derived in the last chapter are not articulated enough if number of agents is too small.



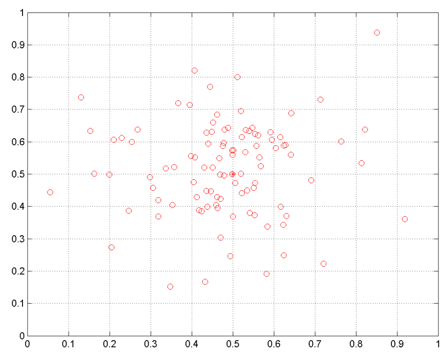
(a) 31 s



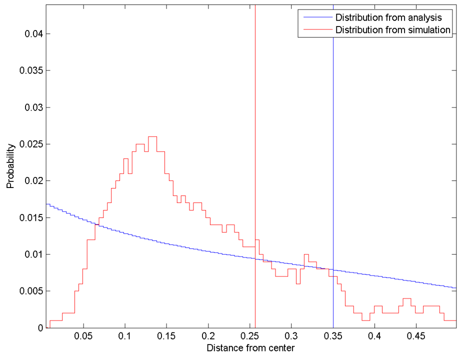
(b) 31 s



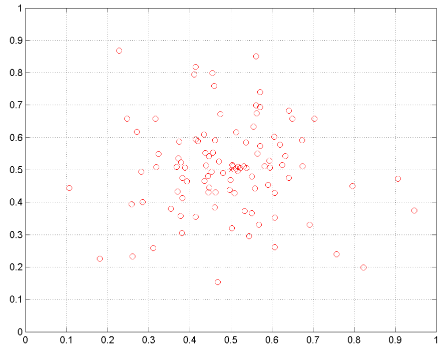
(c) 171 s



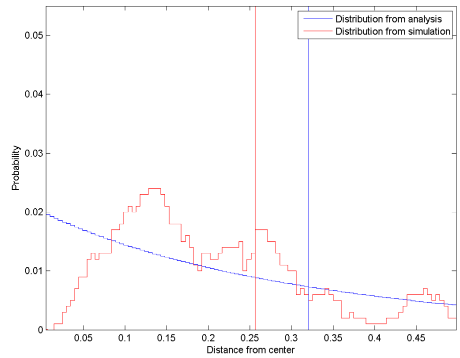
(d) 171 s



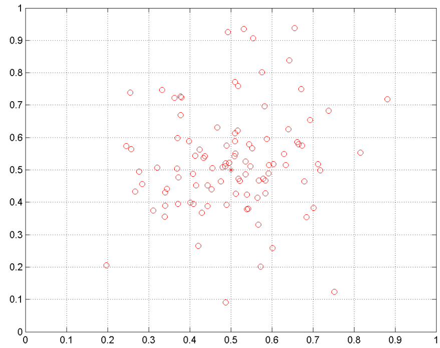
(e) 246 s



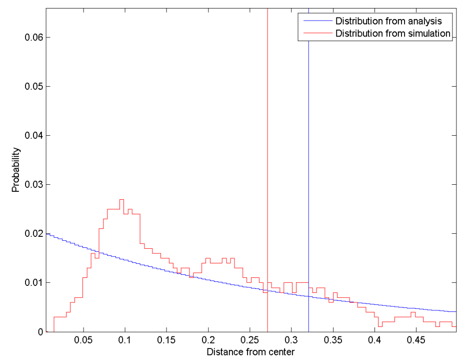
(f) 246 s



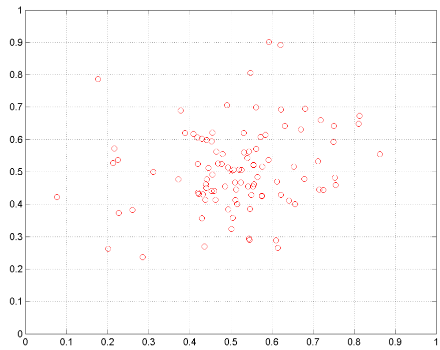
(g) 476 s



(h) 476 s



(i) 590 s



(j) 590 s

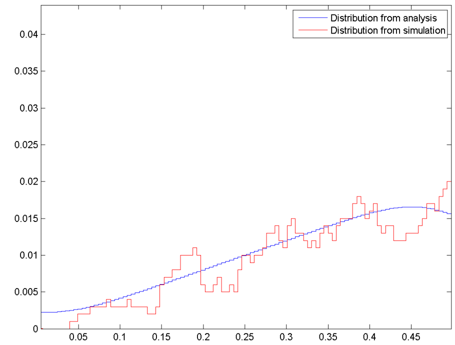
Figure 6.1: Probability distribution prediction vs. distribution of number from simulation

6.1. α

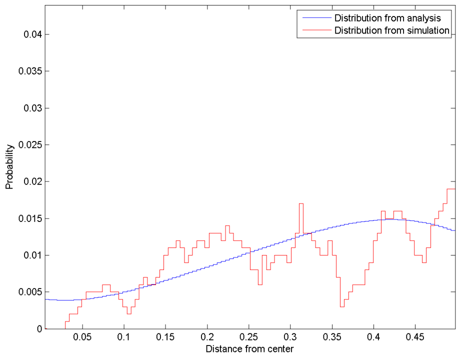
α affects the randomness of movement or stated differently it models agent's tendency of moving towards the heat center. Smaller values of α will make movement more random and therefore, the event of aggregation would be less probable. Both system simulation and probabilistic analysis confirm that (compare Figures 6.2, 6.3 and 6.4)

There are values of α that are too small to produce any inclination towards the cluster formation (Figure 6.4). Again, probabilistic analysis and system simulation yield similar results which verifies the validity of the theory from the last chapter.

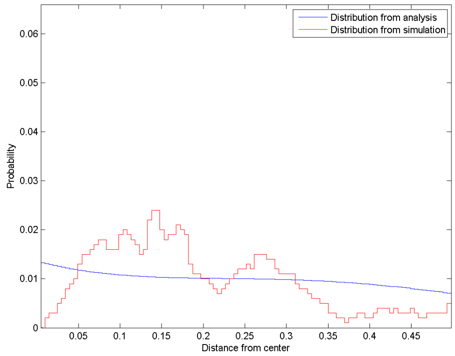
In case of larger α values, when cluster existence is certain and its integrity is strong, the probability analysis detects point of most dense gathering at a distance from the center different from the one that system simulation does. This is obvious on Figure (6.3). However, there is an explanation for this inconsistency. Approximation of circles as straight lines from the last chapter does not hold for circles close to the center. Probability analysis assumes that disks close to center are also straight lines. This assumption produces probabilities for transition towards the center that are greater than the correct ones. Naturally, the probability distribution has its peak at the heat center. Imprecision still does not severely affect the accuracy of cluster prediction, especially the moment of its formation. That inconsistency is expected and it does not affect the overall validity of the analysis from the last chapter. Inaccuracy can be simply treated by taking real geometry into account when calculating probabilities of transition.



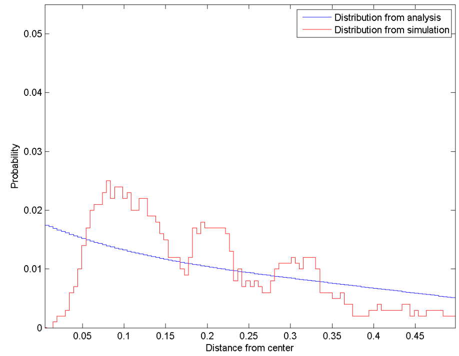
(a) 15 s



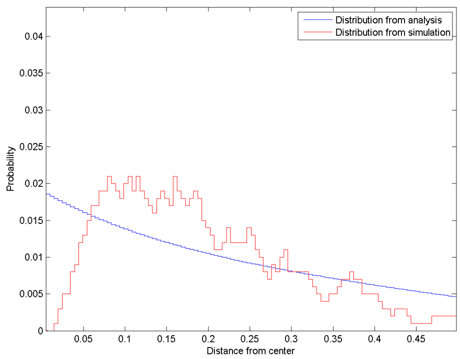
(b) 25 s



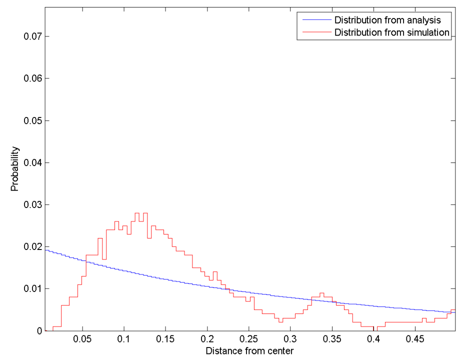
(c) 150 s



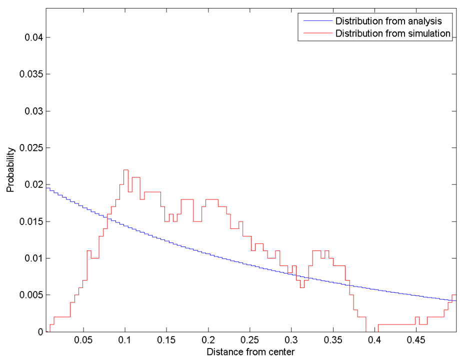
(d) 275 s



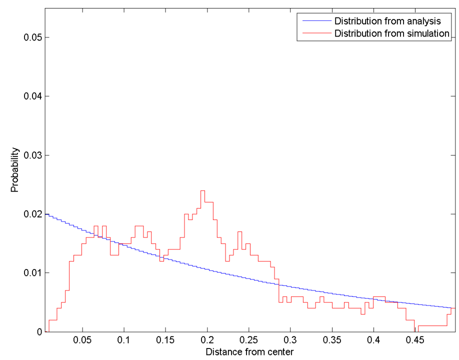
(e) 345 s



(f) 415 s

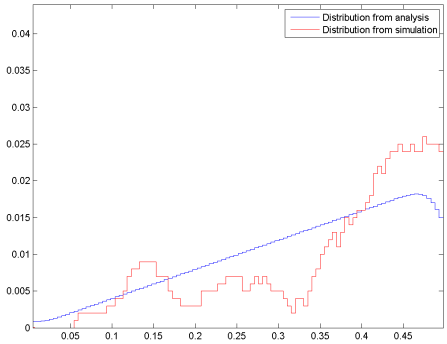


(g) 470 s

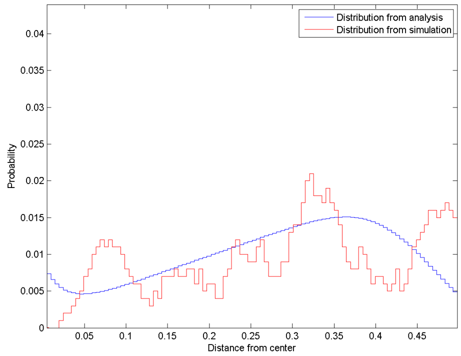


(h) 655 s

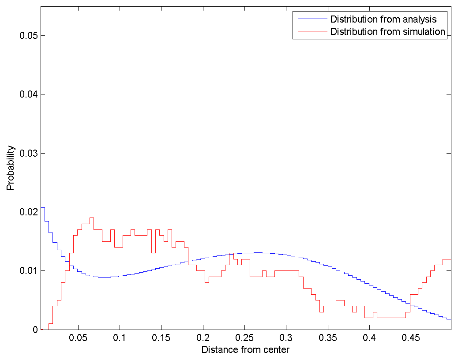
Figure 6.2: Simulation with $\alpha = 0.1$



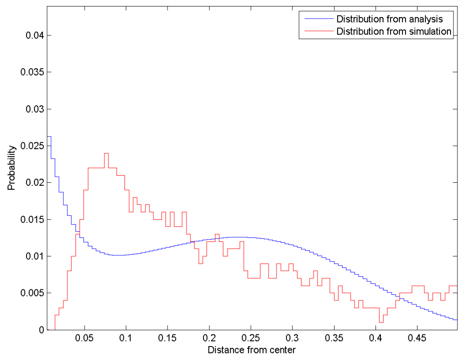
(a) 1 s



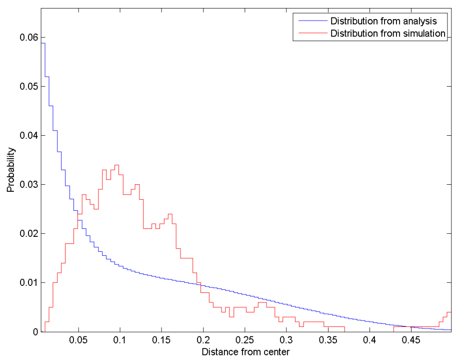
(b) 11 s



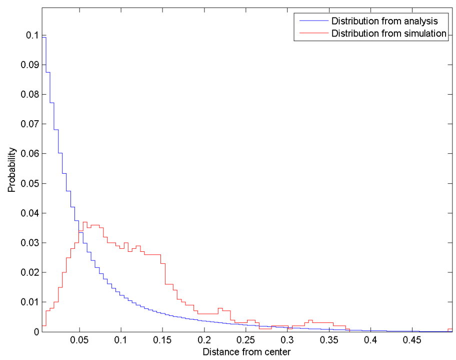
(c) 26 s



(d) 31 s

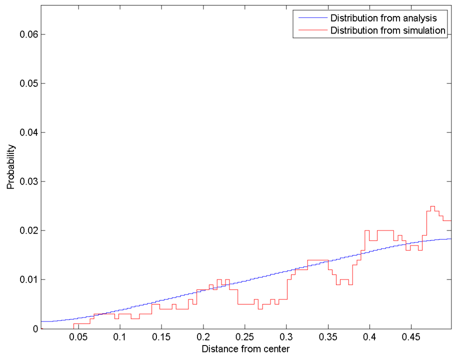


(e) 51 s

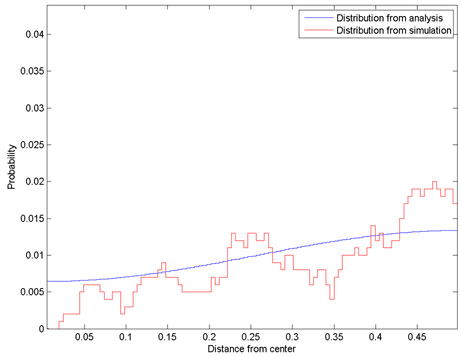


(f) 96 s

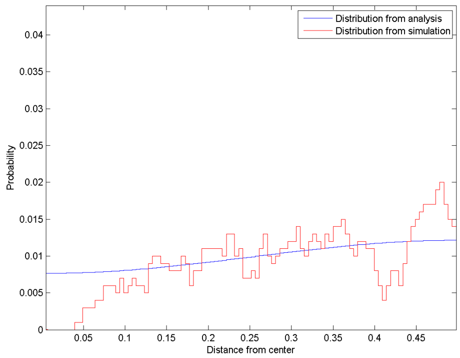
Figure 6.3: Simulation with $\alpha = 0.15$



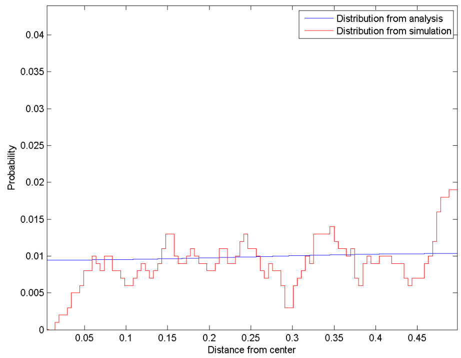
(a) 6 s



(b) 116 s



(c) 176 s



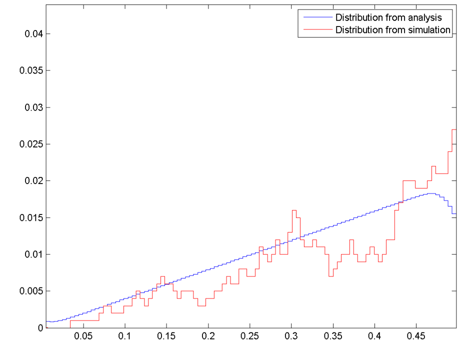
(d) 396 s

Figure 6.4: Simulation with $\alpha = 0.05$

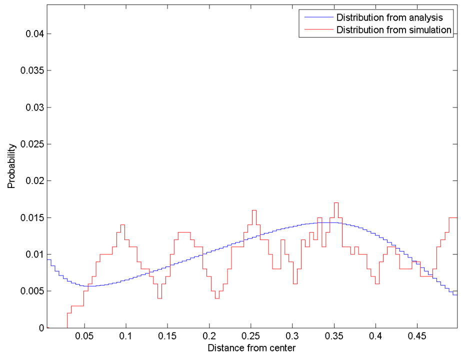
6.2. Random Angle Range

Second parameter that was tested is the deviation of random angle from expected value. Alterations of this parameter are similar to the alterations of α.

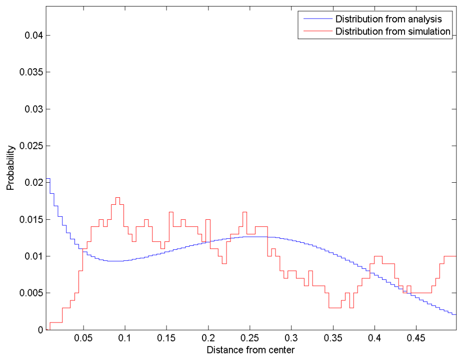
The greater the range, the more freedom is given to the random choice of direction. Figure 6.5 shows how reducing the range by 20% of the previous range affects the speed and intensity of aggregation (look at 6.2 for comparison). As expected, colony was more directed towards the heat center which caused faster aggregation process.



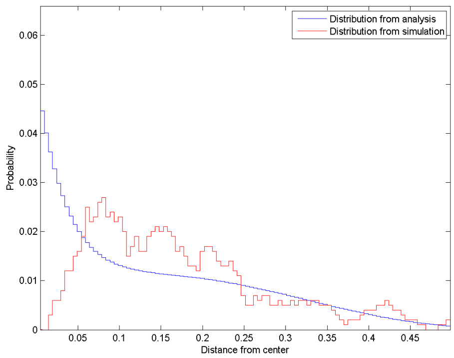
(a) 1 s



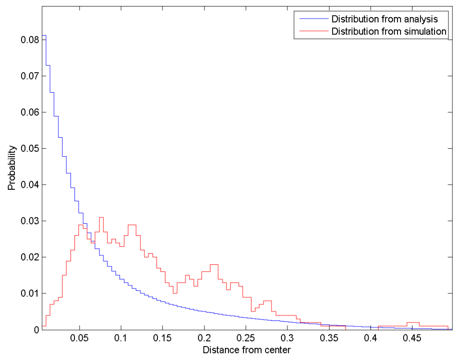
(b) 16 s



(c) 31 s



(d) 56 s



(e) 102 s

Figure 6.5: Simulation with random angle range $[-0.4 \cdot \pi, 0.4 \cdot \pi]$

6.3. Waiting Time and Number of Agents

The following section shows how waiting time and number of agents affect distribution of agents. Previous chapter elaborates how probabilities are calculated when waiting time is considered in model, but theoretical results have to be verified by simulation.

6.3.1. Waiting Time

Two different functions for mapping temperature to waiting time are used in simulation and prediction as shown on Figure 6.6.

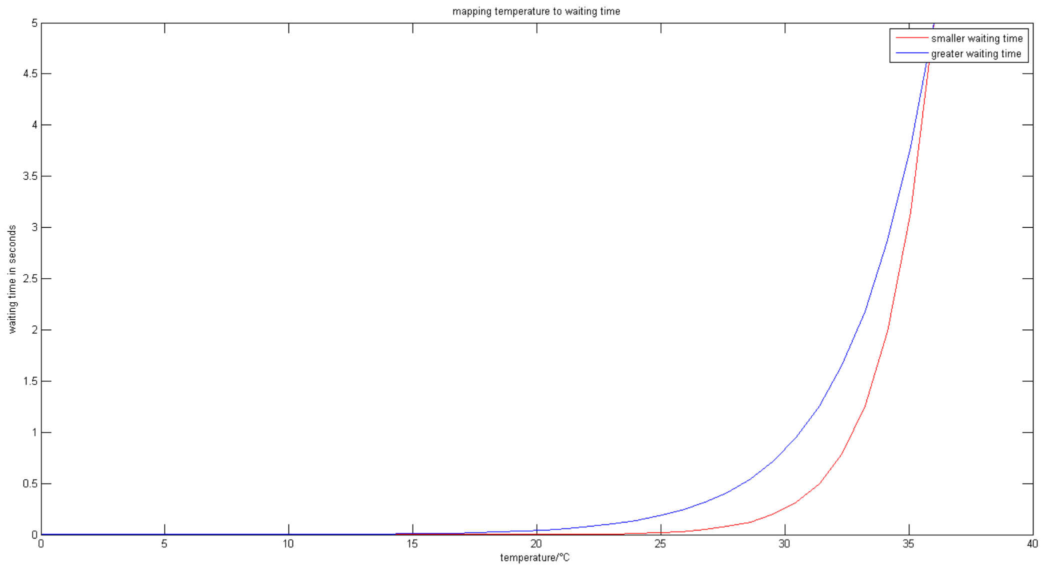
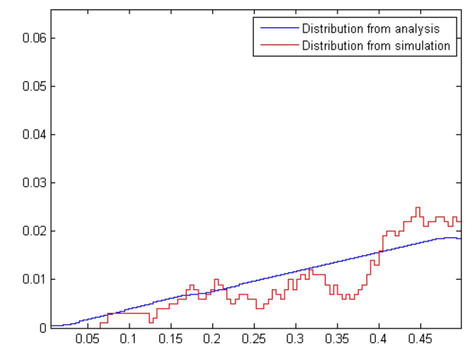
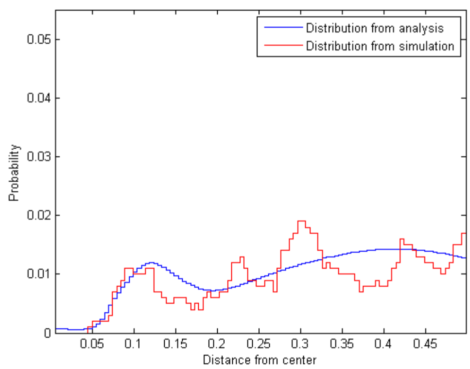


Figure 6.6: Waiting time functions used in simulation

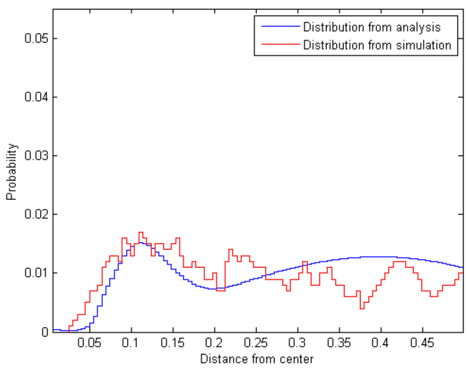
Areas with larger waiting time produce more coherent cluster slightly faster than the areas with lesser waiting time (Figures 6.7 and 6.8). Besides, it can be noticed that predictions with small waiting time are not aggressive enough and bring larger inaccuracies compared to the systems with larger waiting time.



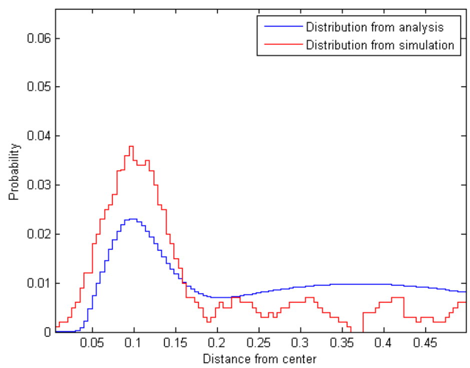
(a) 1 s



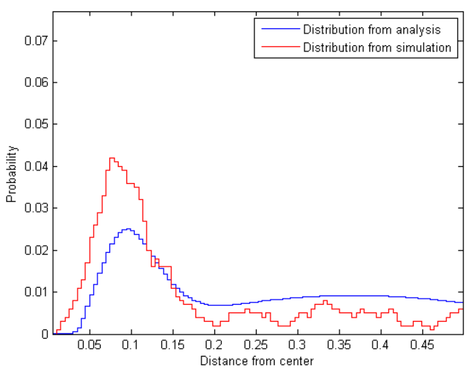
(b) 31 s



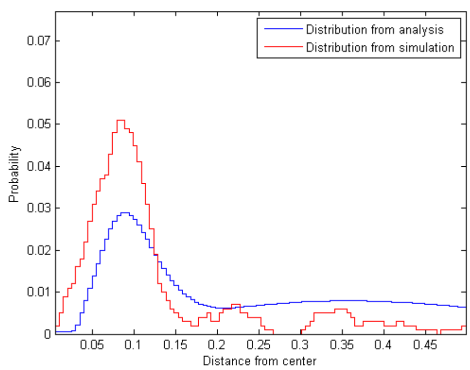
(c) 51 s



(d) 106 s

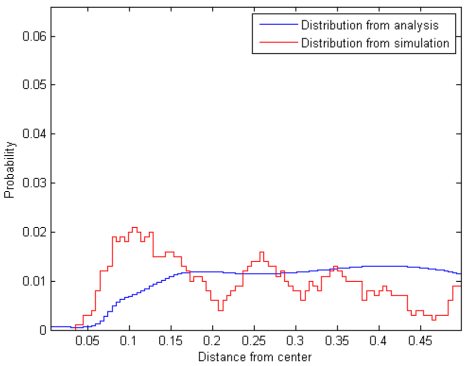


(e) 121 s

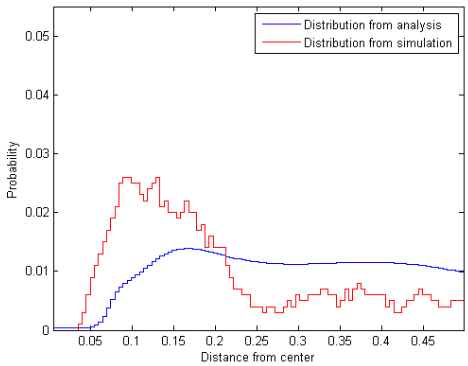


(f) 156 s

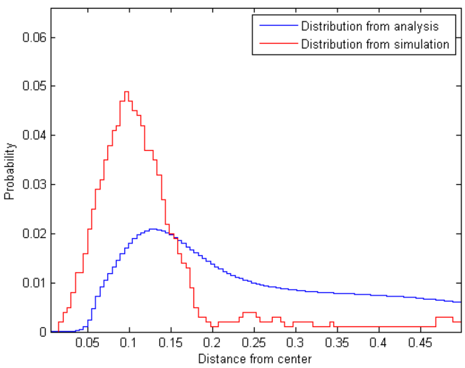
Figure 6.7: Simulation with larger waiting time



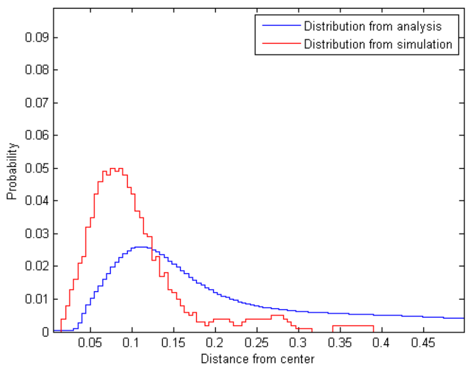
(a) 56 s



(b) 86 s



(c) 206 s



(d) 331 s

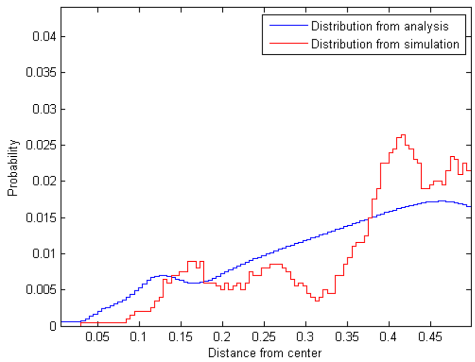
Figure 6.8: Simulation with smaller waiting time

6.3.2. Number of Agents

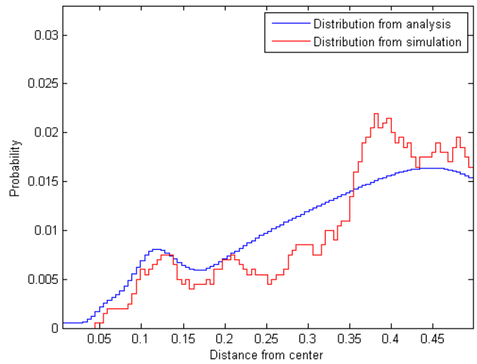
Changing number of agents from 50 to 200 did not significantly influenced simulation nor prediction. Cluster maximum is slightly shifted towards the center in case of smaller number of agents. Greater number of agents start to crate jam around more distant circle than smaller number of agent does. Due to the smaller number, agents are able to move freely when they are close to center until they also reach their jamming point and form aggregation maximum there.

Another thing to notice is that greater number of agent produce behaviour that is more in accordance with stochastic theory. The fact that number of agents affect the quality of stochastic analysis were mentioned few times in thesis and this Figures now prove that.

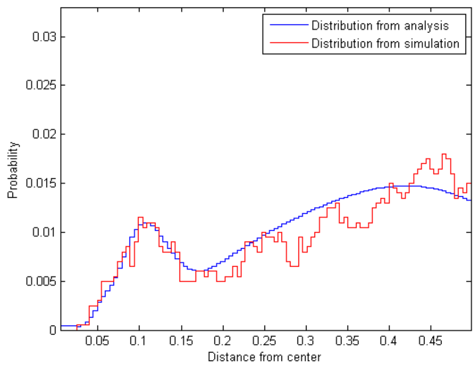
All in all, position of cluster maximum was predicted quite accurately and significant perturbation of number of agents did not affect prediction significantly.



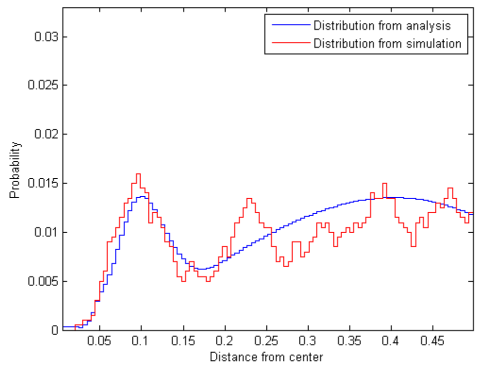
(a) 6 s



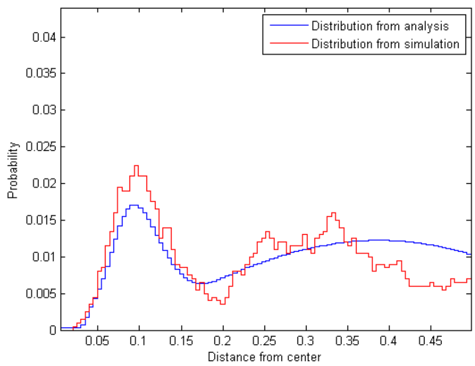
(b) 11 s



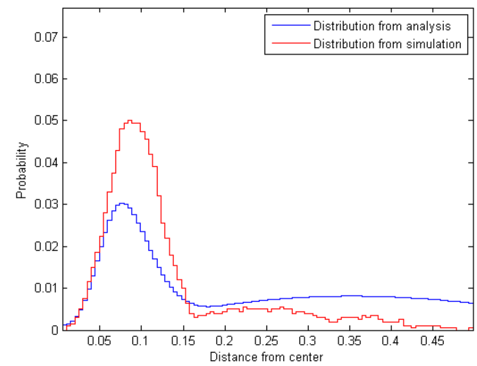
(c) 26 s



(d) 41 s

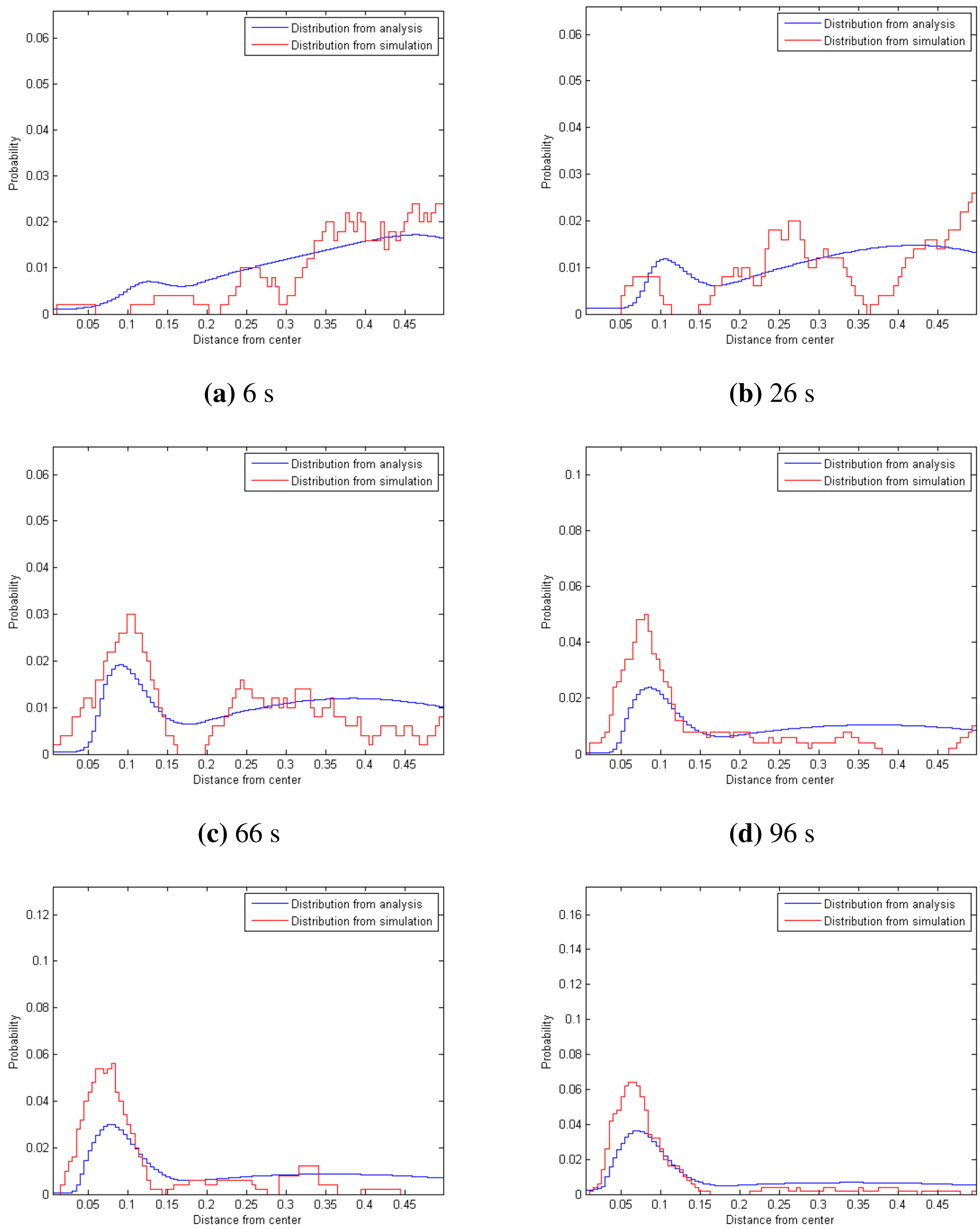


(e) 61 s



(f) 156 s

Figure 6.9: Simulation with 200 agents



(a) 6 s **(b)** 26 s

(c) 66 s **(d)** 96 s

(e) 141 s **(f)** 206 s

Figure 6.10: Simulation with 50 agents

7. Conclusion

Hybrid model of an agent with two discrete states, each modeled by distinct system's equations, produced colony behaviour that was compared with behaviour of bees in the real arena. Simulation and real experiments showed similar dynamics of colony formation under influence of heat source at the center of the arena. Important parameters of the model that were tested are: randomness of agent movement, inclination towards the center of heat, velocity, waiting time, number of agents, etc. They influence the speed of cluster formation and it's integrity.

Once again, Discrete Time Markov Chains proved to be an excellent tool for modeling biological systems' dynamics. Slow process of colony formation enables relatively large unit of time discretization which in turn makes DTMC applicable for the analysis. Convenience of DTMC is that the whole problem amounts to the calculation of state transition probabilities. Although the task of finding these probabilities often turns out to be involving, it was still manageable in the case of this thesis. Few approximation that were made to ease the calculation of probabilities slightly affected the accuracy of prediction, but did not dispute the general validity of theoretical results.
Parameters that describe single agent's dynamics were successfully abstracted in the form of spatial movement probabilities and were encoded into generic macro-level mathematical object, Transition matrix. Transition matrix together with initial distribution of agents in space can yield information about the evolution of spatial distribution vector in time. Spatial movement probabilities were calculated as probabilities of moving closer or farther from the center of the heat. Since the problem is symmetrical it was practical to introduce disks around the heat center. Those disks presented the states of Markov Chains model.

Simulation and evolution of probability distribution generated by Transition Matrix

were highly correlated. Probability distribution followed simulation in both, space and time domain, except for sparse inaccuracies due to small number of agents (small in terms of stochastic theory).
Capability of analysis to detect cluster is even more robust if cluster is defined as a percentage of agents held in a certain radius around heat center. Probability distribution from Markov Chain prediction quite accurately detects the percentage of agents enclosed by an arbitrary circle around heat center (error is less that 10%). Alterations of agent parameters also affected colony distribution as expected, which was additional confirmation of precision.

Thesis was attempt to explore if Discrete Time Markov Chains are suitable for the task and considering all the approximation that were made and results that were still fairly satisfactory, it turns out that DTMC truly is suitable. However, without all the approximations, analysis might have been even more precise. Taking real geometry into account and investigating arena with more than one center might be the next steps.

BIBLIOGRAPHY

Linda J. S. Allen. *An Introduction to Stochastic Processes with Applications to Biology*. 2010.

Neven Elezović. *Vjerojatnost i Statistika*. 2007.

Karlo Griparic. Razvoj bio-hibridnog kolektivnog adaptivnog sustava.

James Hereford. *Analysis of BEECLUST swarm algorithm*. 2011.

R. Thenius Gerald Radspieler K. Crailsheim M. Szopek*, T. Schmickl. Dynamics of collective decision making of honeybees in complex temperature fields. *PLOS ONE*.

Martina Szopek Thomas Schmickl Michael Bodi, Ronald Thenius. Interaction of robot swarms using the honeybee-inspired control algorithm beeclust. *Mathematical and Computer Modelling of Dynamical Systems, Taylor Francis*, 18(1):87–100, 2012.

S. Kernbach T. Schmickl, S. Bogdan. Assisi: Charged hot bees shakin' in the spotlight. *7th International Conference on Self-Adaptive and Self-Organizing Systems*, 2013.

Heinz Worn Karl Crailsheim Thomas Schmickl, Heiko Hamann. Two different approaches to a macroscopic model of a bio-inspired robotic swarm. *Elsevier*, 57: 913–921, 2009.

Title

Abstract

Thesis considers development of agent model that moves in 2D space. Te model should be based on stochastic hybrid automaton. The model parameters should be determined based on honeybee movement in closed 2D space. Influence of the model parameters (caused by changes in the environment) on the aggregation of multi-agent system should be explored. Validity of the model should be tested by simulation in Matlab and compared with honeybee behaviour. **Keywords:** Discrete Time Markov

Chain, Hybrid system, multi-agent system, bio-inspired algorithm, beeclust algoritam

Agent model based on hybrid stochastic automata

Sažetak

Radom na zadatku potrebno je izraditi model gibanja agenta u dvodimenzionalnom prostoru. Model treba biti zasnovan na stohastičkom hibridnom automatu. Parametre modela potrebno je identificirati iz hoda pčele u zatvorenom prostoru. Nadalje, potrebno je ispitati utjecaj parametara na agregaciju više agenata, uzrokovanu promjenama u okolini (zagrijavanje određenog dijela dvodimenzionalnog prostora). Valjanost modela potrebno je ispitati simulacijom u Matlabu uz različite vrijednosti identificiranih parametara i usporedbom sa snimkama hoda pčele.

Ključne riječi: Vremenski diskretni Markovljevi Lanci, Hibridni sustav, multi-agentski sustav, bio-inspirirani algoritmi, beeclust algoritam