

FOOD FOR THOUGHT:  
WHEN INFORMATION OPTIMIZATION FAILS TO OPTIMIZE UTILITY

by

EDWARD K. AGARWALA

Submitted in partial fulfillment of the requirements

For the degree of Master of Science

Thesis Adviser: Dr. Peter J. Thomas

Department of Mathematics

CASE WESTERN RESERVE UNIVERSITY

August, 2009

# Contents

<b>List of Tables</b>	<b>iv</b>
<b>List of Figures</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vi</b>
<b>Abstract</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Information Theory . . . . .	2
1.2 Model . . . . .	4
1.2.1 Strategies . . . . .	4
1.2.2 Distance Certain Model . . . . .	5
1.3 Results . . . . .	5
<b>2 Model</b>	<b>6</b>
2.1 Overview . . . . .	6
2.1.1 World . . . . .	6
2.1.2 Food Distribution . . . . .	7
2.1.3 Creatures . . . . .	8
2.1.4 Energy . . . . .	9
2.1.5 Event Sequence . . . . .	10
2.2 Formal Model . . . . .	11
2.2.1 Time . . . . .	11
2.2.2 World . . . . .	12
2.2.3 Injector . . . . .	12
2.2.4 Molecules . . . . .	14
2.2.5 Creatures . . . . .	15
2.2.6 Event Sequence . . . . .	18

<b>3</b>	<b>Molecules in the World</b>	<b>22</b>
3.1	Overview . . . . .	22
3.2	Decay/Movement with Poisson Random Variables . . . . .	23
3.3	The Sum of Two Poisson Random Variables . . . . .	25
3.4	Molecule Counts are Poisson . . . . .	25
3.5	Calculating $\mathbb{E}(K_i)$ ( $\lambda_i$ ) . . . . .	27
3.6	Properties of the Mean Concentration Profile $\{\lambda_i\}_{i=0}^{N-1}$ . . . . .	29
3.7	The Effect on $\{K_i\}_{i=0}^{N-1}$ as $\gamma \rightarrow \infty$ . . . . .	34
<b>4</b>	<b>Using Sense Data</b>	<b>36</b>
4.1	Overview . . . . .	36
4.2	Review of Notation . . . . .	37
4.3	The Injector's Current Location, Given a Full History . . . . .	38
4.4	The Injector's Next Location, Given a Full History . . . . .	40
<b>5</b>	<b>Injection Rates in the Limit</b>	<b>42</b>
<b>6</b>	<b>Coarse Graining for a Simplified Model</b>	<b>48</b>
6.1	Introduction . . . . .	48
6.2	Information States for $F_+^t$ . . . . .	48
6.3	Information States for $F_-^t$ . . . . .	50
6.3.1	The Information State $\mathcal{J}_-^0$ . . . . .	52
6.3.2	The Information State $\mathcal{J}_-^1$ . . . . .	55
6.3.3	The Information State $\mathcal{J}_-^2$ . . . . .	58
6.3.4	The Information State $\mathcal{J}_-^{3+}$ . . . . .	61
6.3.5	The Information State $\mathcal{J}_-^I$ . . . . .	62
6.3.6	Entropy Minimal State . . . . .	65
<b>7</b>	<b>Creature Strategies</b>	<b>67</b>
7.1	Information . . . . .	68

7.2	Information Theory Creature . . . . .	69
7.3	Maximum Likelihood Creature . . . . .	70
7.4	Modified Maximum Likelihood Creature . . . . .	72
7.5	Results . . . . .	75
<b>8</b>	<b>Discussion</b>	<b>80</b>
8.1	Alternative Food Distributions . . . . .	80
8.2	The Injector's Movement Algorithm . . . . .	80
8.3	Survival Time . . . . .	81
<b>9</b>	<b>Conclusion</b>	<b>82</b>

## List of Tables

2.1	Standard Parameter Values . . . . .	19
2.2	Random Variables . . . . .	20
2.3	Variables and Notation . . . . .	21

# List of Figures

3.1	Poisson Random Variable Means $\lambda'_i$ s. . . . .	35
5.1	A Diagram of Locations and Distances for Theorem 5.1. . . . .	43
6.1	The Support of $F_+^t$ in the Distance-Certain Model. . . . .	51
6.2	A Probability Mass Function in $\mathcal{J}_-^0$ . . . . .	53
6.3	Transitions from $\mathcal{J}_-^0$ . . . . .	53
6.4	A Probability Mass Function in $\mathcal{J}_-^1$ . . . . .	56
6.5	Transitions from $\mathcal{J}_-^1$ . . . . .	56
6.6	A Probability Mass Function in $\mathcal{J}_-^2$ . . . . .	59
6.7	Transitions from $\mathcal{J}_-^2$ . . . . .	59
6.8	A Probability Mass Function in $\mathcal{J}_-^{3+}$ . . . . .	63
6.9	Transitions from $\mathcal{J}_-^{3+}$ . . . . .	63
7.1	Information Theory Creature State Transitions. . . . .	71
7.2	Maximum Likelihood Creature State Transitions. . . . .	73
7.3	Modified Maximum Likelihood Creature State Transitions. . . . .	76
7.4	Information for Each Strategy Given $P_x$ . . . . .	79
7.5	Expected Food for Each Strategy Given $P_x$ . . . . .	79

## Acknowledgements

I am profoundly grateful to my thesis committee Dr. Hillel Chiel, Dr. Ken Loparo, Elizabeth Meckes, and my advisor Dr. Peter Thomas for their extraordinary help in developing and editing this work. Each member of the committee has dedicated time and effort above and beyond the call of duty.

This document could not have been written without the enduring love and support of my family and friends. I thank Dr. Vijaya Agarwala (father), Dr. Kathleen Kane (mother), Matthew Agarwala (brother), Mary Agarwala (sister), Kate Moore and Catherine Vermeersch; neither my work nor I would be here without their aid.

Food For Thought:  
When Information Optimization Fails to Optimize Utility

Abstract

by

EDWARD K. AGARWALA

Information maximization criteria have been used to account for the physiology of sensory systems as diverse as receptive fields in the primary visual and auditory cortices, and olfaction. We investigated a model of an organism searching for food by taking successive samples from an environment in which food particles diffuse stochastically from a slowly and randomly moving source. In the limit of large food concentrations we reduced our high dimensional model system to a Markov chain on a small number of equivalence classes. In this system we made rigorous quantitative comparisons of different search strategies based on (i) maximizing the searcher's information about the food source's location, (ii) maximizing the likelihood of landing on the source, and (iii) hybrid strategies combining aspects of (i) and (ii). In terms of long-term expected food benefit we found that each strategy was superior to the others depending on the source's rate of movement.



# 1 Introduction

Information theory provides a mathematical framework for understanding communication systems [10, 11]. The success of information theory lies in its abstraction away from notions of semantics, or relative importance of different messages. A line of research has developed in recent years that can account for some features of the physiology of sensory systems in terms of principles of information maximization [2, 3, 8, 12, 13]. There has also been great interest in applying the mathematical theory of communication to the genetic code [4–6], *c.f.* C. Adami’s statement that “the discovery of the genetic code cemented the fact that information is the *central* pillar in any attempt to understand life” [1, pg. 59, original emphasis]. However, the very abstraction of Shannon’s framework poses a stumbling block for attempts at the principled application of information theory to biology. It is a commonplace that in biological systems, “information” is always *about something*. Appropriate selection of which information to retain and which to ignore can significantly impact an organism’s fitness and survival. But despite attempts to extend information theory to encompass a subjective element [7] there is little consensus in the field on the way forward.

This thesis is part of a longer range research program to devise a series of model systems in which information from multiple sources is selectively enhanced or ignored according to the needs of an organism. We wish to make the model precise enough to establish unambiguously the optimality or non-optimality of different information strategies. Consider, for example, a foraging creature which requires two types of food,  $A$  and  $B$ , available from two independent randomly located sources. Depending on the metabolic requirements and reserves of the two foodstuffs, slightly increasing the creature’s information about one at the cost of greatly reducing information about the other might be the best survival strategy, even though it leads to a net reduction in total information.

As a first step to making such a framework precise enough for careful analysis,

we consider in this thesis a single creature and compare foraging strategies based on maximizing information about the source *versus* maximizing the expected food intake from the source. We will make a number of simplifying assumptions, the strongest of which is that the model creature has *perfect information* about all aspects of the food source *excepting only its location*. The food propagates *via* rapid diffusion from a source that can itself diffuse, albeit slowly. Given observations of the quantity of food present at a succession of locations, the “creature” can try to infer the location of the source and move accordingly, *via* discrete time steps. Should it move in such a way as to maximize its information about the location of the food source? Could this strategy possibly have different consequences than merely trying to move to the creature’s current best guess as to the source location? Even with a single creature and a single food source we find that, remarkably, the two strategies are not identical. In particular, under some conditions *there is no information optimal strategy which is also a food optimal strategy*. This observation may seem obvious in retrospect but perhaps commands more interest on account of the computational neurosciences’ long standing acceptance of information maximization as an explanatory principle.

## 1.1 Information Theory

Information theory originated as the study of how much information can be transmitted through a physical medium such as a telegraph wire. Since Shannon’s founding treatise [10, 11], the tools and framework he introduced have allowed us to ask quite general questions such as: *How much information can be gleaned about the random results of one experiment given the random outcome of another experiment?* For example, how much can one tell about the weather given one sees someone wearing shorts? How much can a creature know about its environment given what it has observed of its immediate surroundings? The use of the term “how much” implies that there is quantitative measure of information. The *mutual information* between two random variables (defined below) quantifies how much one random variable is

statistically related to another.

*Entropy* quantifies the uncertainty regarding the random outcome of an experiment [10, 11]. The entropy of a discrete random variable  $X$  is

$$H(X) := - \sum_{\{x:\mathbb{P}(X=x)\neq 0\}} \mathbb{P}(X = x) \log_2(\mathbb{P}(X = x)). \quad (1.1.1)$$

Entropy is measured in bits and in particular  $H(X)$  is the minimum number of bits required (on average) to report a realization of  $X$ . One important property of  $H(X)$  is that it is non-negative for any random variable  $X$ .

The *conditional entropy* of  $X$  given  $Y$  is

$$\begin{aligned} H(X|Y) &:= \mathbb{E}_Y(H(X|Y = y)) \\ &= \sum_y \mathbb{P}(Y = y) H(X|Y = y). \end{aligned} \quad (1.1.2)$$

This represents the expected reduction in uncertainty about a realization of  $X$  given a realization of  $Y$ . It is always true that  $H(X|Y) \leq H(X)$ . This means the uncertainty of  $X$  cannot on average be increased by the random variable  $Y$ ; however, for some  $y$   $H(X|Y = y)$  may be greater than  $H(X)$ .  $H(X)$  is a bound on the average  $H(X|Y = y)$  for all  $y$ .

The mutual information between two random variables  $X$  and  $Y$  is

$$H(X; Y) := H(X) - H(X|Y). \quad (1.1.3)$$

The mutual information also quantifies the departure from independence of the random variables  $X$  and  $Y$ . A high mutual information indicates that the variables are dependent upon each other. Two random variables are independent if and only if they have 0 bits of mutual information.

An analogy can be drawn to the correlation between two variables. A high correlation coefficient means the random variables are linearly dependent (statistically if not causally). The primary reason the analogy fails is that information theory removes the semantics from the question and asks solely whether the random variables

are statistically dependent. Consider the random variables  $X$  where  $\mathbb{P}(X = 1) = 1/2$  and  $\mathbb{P}(X = 2) = 1/2$  and  $Y$  where  $\mathbb{P}(Y = X) = 1/2$  and  $\mathbb{P}(Y = -X) = 1/2$ . In this case the correlation coefficient between  $X$  and  $Y$  is 0 but the mutual information between the two is 1 bit. In general, if two random variables are correlated they are statistically dependent but not *vice versa*.

## 1.2 Model

We develop a simple foraging model to test the effectiveness of different heuristic behaviors. We have a discrete circular world where a creature, an injector and molecules reside. The injector injects molecules into the world that can move about and decay. For certain rules on this behavior the food at each location on the ring is distributed according to its distance from the injector. The creature can observe these molecules and use its past observations to predict the injector's next location. How a creature responds to this information depends upon its strategy.

### 1.2.1 Strategies

The strategies that will be reviewed in this thesis include an information optimal strategy, a maximum likelihood strategy, and a hybrid of the two. Although these strategies can be named it takes work to show that a creature's response to the environment lives up to its strategy's name. These strategies will be properly defined in Section 7. For each strategy we obtain we will calculate the average amount of food seen and the average amount of information received per time step. It is on this basis that strategies will be compared.

As an overview, the information optimal strategy will maximize the creature's knowledge about food source's location in the world. The maximum likelihood strategy will always go where it thinks the injector will be. The modified maximum likelihood will be a hybrid of the two.

### 1.2.2 Distance Certain Model

Based on work describing the food distribution and the creature's ability to predict the injector's next location we conjecture that in the case of high injection rates the creature becomes confident of the injector's location. Furthermore, we make the assumption that the model with injection rate going to infinity is a good approximation of the model for high injection rates.

If we accept these conjectures, the creature strategies can be viewed as driving a Markov process. It is possible to calculate the expected food, the mutual information between the creature's history and the injector's location, and the transition probabilities for each Markov state. From there we calculate the expected food and information for each strategy.

## 1.3 Results

This model demonstrates that information maximization does not guarantee utility optimization. For a set of parameters, each of the three strategies does better than the other two in terms of average food, but the information optimal strategy is strictly better than the modified maximum likelihood strategy which is strictly better than the maximum likelihood strategy in terms of information.

## 2 Model

### 2.1 Overview

The system investigated in this thesis has a natural interpretation as a model of foraging behavior. Our goal is to analyze and compare different heuristics for guiding such behavior (information maximization, maximum likelihood estimation, *et cetera*). In order to compare foraging heuristics, there must be some sort of model in which creatures can forage for food. In devising any model eventually one must make (sometimes arbitrary) decisions about how the world will work. When such decisions are made, it is sometimes true that an alternate choice would be equally reasonable and create no major changes in the methodology though perhaps in the results. On the other hand some choices are made to make the problem tractable. In those cases an indication of why the simplification is useful but still reasonable shall be made. We will discuss these alternatives informally at the beginning of relevant sections.

#### 2.1.1 World

Our model observes one creature foraging for one kind of food in a one-dimensional world. Only one creature is observed because the goal is to compare foraging rather than competitive-foraging heuristics. Therefore this thesis is primarily applicable to essentially solitary foragers and foragers that are not in strong competition. This could occur when the amount of food available is not significantly impacted by the number of foragers in the area. For example, one could imagine creatures feeding on a large, dense swarm of flies. In this case the creatures may not significantly impact the total number of flies but each creature might get significantly more or less flies depending on its strategy.

There is only one kind of food because all creatures need at least one source of food and this thesis can be a foundation for models with more types of food. The world is one-dimensional to make analysis tractable. That being said many creatures

do feed along essentially one-dimensional regions; for example, rivers, reefs, the edges of a forest or the shores of a body of water are one-dimensional feeding grounds.

The world is also discrete, finite and circular. The world is discrete so that Shannon's discrete entropy measure can be used. The world is finite and circular primarily to make computer simulations feasible.<sup>1</sup> It also aids in the analysis.

### 2.1.2 Food Distribution

In order to meaningfully test the heuristics, the distribution of food needs to have certain characteristics. Firstly, there needs to be a finite amount of food distributed throughout the world such that the food is not evenly distributed. If there is an equal amount of food everywhere there is no need to search for "better" locations to feed. Secondly, there must be some order to how the food is distributed so that there is a way to search for "better" locations to feed. In this case, "better" could mean the location with the most food or the location which on average will allow the creature to obtain the most food, but for the purposes of this discussion "better" should be related to creature's foraging heuristic.<sup>2</sup>

In this model, there is a single (possibly mobile) injector or source that introduces molecules (units of food) into the world. The molecules can then move about the world and possibly decay. For certain algorithms of movement, decay and injection, a unimodal distribution of food occurs and the creature could climb the gradient to "better" feeding grounds. There are other examples of food distribution that meet the two criteria from above. Many organisms swarm around places with good food or shelter and that good location will have a higher density of organisms than nearby. As a concrete example consider the densities of birds or fish when they are given bread. Any creature foraging for those organisms would be feeding on a unimodal distribution of food with a gradient.

---

<sup>1</sup>Simulations are planned for future work.

<sup>2</sup>Information Optimal, Maximum Likelihood or modified Maximum Likelihood

One important consideration for this model is that there is a single injector of food. A single injector simplifies analysis and is a reasonable approximator for the situation when injectors are far apart. Consider a situation where food from one source is highly unlikely to reach the other.

### 2.1.3 Creatures

The creatures in this model are designed to be the best possible creature that could exist in the system. They have full memory of everywhere they have been and the amount of food at those locations. They know how the world works in terms of molecule decay, movement and injection. Furthermore, the creatures can quickly move to any location in the world. These strong assumptions are made to create an upper-bound on what creatures could possibly do. It turns out these faculties are equivalent to much weaker abilities such as maintaining a state of where the injector (the center of food distribution) is likely to be, knowing the steady state distribution of molecules and moving only a few spaces quickly as opposed to any where in the world. This is discussed further in Section 4.

In this thesis, it is often stated that the creature “knows” something about the world. This not to say the creature knows some fact cognitively but that it reacts as if it did. For example, a child may catch a ball or frisbee without cognitively knowing the vector calculus involved in predicting the flight path. The specific mechanism that allows the child to figure out the flight path is separable from the observable behavior, namely the child can consistently catch the ball. This thesis is concerned with the best possible behaviors that satisfy different heuristics rather than attempting to find or devise a biological mechanism that does the job.

The creature knows two more facts: how the injector moves and how the creature and injector start in the world. There are many ways the creature and injector could start in the world. They could both start at the same location similar to when a creature that is born near a good source of food. Alternatively, the creature could



know nothing about the injector's starting location other than there must be one. This is equivalent to knowing the injector is at a random location (uniform distribution). For this thesis, the latter starting condition is chosen.

#### **2.1.4 Energy**

All the foraging heuristics considered in this thesis are compared using the average amount of food observed. This is done because in the long term the average amount of food the creature can get is a reasonable predictor of success and this formulation of question is amenable to analytic techniques. There are of course other considerations: Firstly, there is a question of how the amount of food consumed is dependent on the amount of food seen. Secondly, the average cost in energy of movement can affect the rankings of different heuristics. Finally, there is the question of whether or not the average energy gained and expended is a useful measure of success.

In the framework of this model the creature did not explicitly consume the food. The primary reason for this is that a general approach allows for a wider variety of feeding patterns to be considered. Using the methods outlined in Section 7, it is possible to consider feeding methods where the creature gets all the molecules or has a 50% chance of getting each molecule or 1 molecule as long as there is a molecule to be had. In the first two cases, the average food consumed is proportional to the amount of food seen while the latter is not. For our results (though not our methods) it is assumed that the expected food consumption by the creature is proportional to the amount of food the creature sees.

The creature may jump in a single time step to any location in the world. In the framework of this model the creature does not expend any energy to make these moves. Again, this approach still allows for a calculation of the frequency with which the creature moves each distance. It is assumed that a creature will move to its new location along the shortest path. For example, the creature will not move around the entire world to stay in the same location.

In the real world, creatures move about, feed, breed and die. Survival is often about staying above a certain threshold. Being able to breed is often about obtaining enough energy to survive and produce offspring. In the cases where death is likely and reproduction is rare, the actual amount of energy the creature has at any given time is significantly more important than the average amount of energy. The analytic methods used in this thesis are not conducive to comparing different heuristics in these situations. Future work will use computational methods capable of comparing the heuristics but it requires a great deal of specification in terms of all the parameters to run the simulations.

### 2.1.5 Event Sequence

If we want to model a creature and an injector moving in discrete time intervals, it is necessary to have a ratio of creature movements to injector movements. In this thesis we focus our attentions on a ratio of 1:1. This is simply a convenient choice. A mathematical discussion of other ratios can be found in Section 8.2. During a time interval an injector moves to a location, the molecules distribute to steady state and then the creature observes the number of molecules at a location of its choice. This cycle repeats *ad infinitum*.

A primary objection to this methodology is the fact that the creature sees nothing while it is moving its next location nor while the molecules go to steady state. In some scenarios this is entirely reasonable. Consider a fishing bird flying from location to location. While it is flying it gains no information about the number fish in the water below. An alternative is that the creature feeds during the day but then sleeps for the night while the distribution of food moves. A creature that feeds in tidal pools gets a glance at the distribution of food in ocean whenever the tides recede. The food in the ocean moves continuously, but the creature only sees a snapshot of the food distribution per cycle of the tides. Finally, it is possible that the food moves significantly faster than the creature or the injector, which explains how the system

can come to steady state between the creature's observations.

## 2.2 Formal Model

The model has four different components: the world in which everything resides, an injector, the molecules injected by that injector and a creature. In order to represent and compare foraging strategies for the creatures, it is necessary to lay out the properties of each component including how they interact and evolve over time.

### 2.2.1 Time

The first consideration is the nature of time. In this model, time is determined by discrete time steps. There are two time scales: a slow time scale for the creature and injector and a fast time scale for the molecules.

Let  $t \in \mathbb{N}$  enumerate the (slow) time steps for the creature and injector.

Let  $\tau \in \mathbb{N}$  enumerate the (fast) time steps for the molecules within some slower time step  $t$ . What happens during each time step  $t$  and  $\tau$  is defined in Section: 2.2.6. It is important to note that the fast time steps are used solely to explain how the molecules decay, move and come into the world. The only changes in the world between  $\tau$  and  $\tau + 1$  involve molecules.

At time  $(t : \tau) = (0 : 0)$ <sup>3</sup> the world contains an injector and a creature, but no molecules. The world starts with no molecules for purely technical reasons; it is shown in Section: 3.4 that the nature of world is independent of any initial distribution of molecules.

These time steps represent snapshots of the state of events in the world. Each distinct event is not given its own time step simply because such enumeration requires much notation and provides little insight into the model's behavior.

---

<sup>3</sup>This notation can be likened to hour:min time notation.

### 2.2.2 World

The creature, the injector and all the molecules exist on a one-dimensional ring with an even number of discrete spatial locations. Let  $N$  be the number of locations on the ring. See Table: 2.1 (page 19) for standard values of parameters used in the conceptual model.

The use of periodic boundary conditions and a world with an even number of locations is mathematically convenient but conceptually irrelevant. For a sufficiently large  $N$  and for creature strategies that keep the creatures close to the injector, there is no qualitative difference in the results for even and odd  $N$ .

We enumerate each location on the ring from 0 to  $N - 1$  in a clockwise fashion. In what follows we will refer to clockwise and counterclockwise displacements respectively as directions “right” and “left”.

When the creature, injector or molecules move, we will sometimes denote the displacements with regards to the original location. For example, a molecule at location  $l$  may move 1 step to the right. The molecule’s new location is then  $(l + 1) \bmod N$ . For ease of notation, a location “ $l$ ” will always refer to  $l \bmod N$ . Furthermore,  $a$  steps to the right of location  $l$  will always be  $l + a$  and  $a$  steps to the left will always be  $l - a$ .

Because the world is a ring the term “distance” needs to be defined. Given two positions on the ring,  $l_1$  and  $l_2$ , let the distance between them be  $\min\{(l_1 - l_2) \bmod N, (l_2 - l_1) \bmod N\}$ .

Throughout this thesis there will be references to probability mass functions on the set  $\mathbb{Z}_N$ . Let  $\mathcal{P}(\mathbb{Z}_N) = \{f \in \mathbb{R}^N : \sum_{l=0}^{N-1} f(l) = 1 \text{ and } f(l) \geq 0 \text{ for all } l\}$ . The set of probability mass function on  $\mathbb{Z}_N$  is the set  $\mathcal{P}(\mathbb{Z}_N)$ .

### 2.2.3 Injector

Molecules enter the world at the site of the injector. In some scenarios the injector or source will be allowed to move about the world but in others will remain fixed. Let

$R^{t:\tau}$  be a Poisson random variable with mean  $\gamma$ .

$$\mathbb{P}(R^{t:\tau} = r) := \frac{e^{-\gamma}\gamma^r}{r!}$$

$R^{t:\tau}$  molecules are injected at the source's location on the fast time step  $t : \tau$ . See Table 2.1 (page 19) for standard values of model parameters.

Let  $S^t$  be the random variable denoting the source's position at time  $t$ . At time  $t = 0$  the injector is assumed to be in the world at a random location (uniform distribution). This implies  $\mathbb{P}(S^0 = s) = 1/N$  for all  $s \in \{0, \dots, N - 1\}$ .

The injector's movement is always described relative to its current location. An injector movement algorithm  $P$  is a probability mass function on  $\mathbb{Z}_N$  where  $\mathbb{P}(P = m)$  denoted by  $p_m$  is the probability of moving  $m$  steps to the right of the current location. There are three movement algorithms that are of particular interest:

1. Algorithm " $P_x$ ":

$$p_m := \begin{cases} 1 - 2x, & m = 0 \\ x, & m = 1 \\ x, & m = N - 1 \end{cases}$$

In this case the injector moves left or right with probability  $x$  or does not move with probability  $1 - 2x$ . Clearly we must require  $0 \leq x \leq 1/2$ . See Table 2.1 for standard parameters values used for simulation or analysis.

2. Algorithm " $P_0$ ":

$$p_0 := 1$$

This is a special case of algorithm  $P_x$ , for  $x = 0$ . In this case the injector is stationary.

3. Algorithm " $P_{uniform}$ ":

$$\forall m, p_m := 1/N$$

In this case the injector is equally likely to move anywhere.

The algorithms  $P_{uniform}$  and  $P_0$  represent upper and lower bounds (respectively) on the entropy of injector movement algorithms.  $H(P_{uniform}) = \log(N)$  and  $H(P_0) = 0$ .

**Proposition 2.1.** *The probability that the injector is at location  $s$  at time  $t$  is  $1/N$ .*  
 $\mathbb{P}(S^t = s) = 1/N$ .

*Proof.* This follows because  $P$  is independent of the injector's location and  $\forall s \mathbb{P}(S^0 = s) = 1/N$ . Consider the injector starting at some location,  $l$ , moving along some path according to  $P$  and ending up at location  $m$ . It is equally probable that the injector could have started at  $l + a$  because

$$\mathbb{P}(S^0 = l) = \mathbb{P}(S^0 = l + a) = 1/N,$$

moved along the same (equally probable) path, and end up at location  $m + a$ .

This is true of all starting locations and paths so  $\forall a, \forall t$

$$\mathbb{P}(S^t = s) = \mathbb{P}(S^t = s + a).$$

Therefore,  $\forall t, \forall s, \mathbb{P}(S^t = s) = 1/N$ . □

#### 2.2.4 Molecules

Molecules are objects that come into the world by means of injection, move about the world through a stochastic algorithm, and leave the world through a decay process. The molecules affect the creature in two different ways; first, they are a source of food and second, the number of molecules the creature sees at a given location can be used to infer where the injector is and hence where more food may be.

Once injected into the world in a particular location, each molecule moves according to a stochastic algorithm independent of the presence of other molecules, the location of the creature, and any subsequent movement of the injector. We describe a molecule's movement relative to its current location in terms of a movement algorithm  $Q$ .  $Q$  is a probability mass function on  $\mathbb{Z}_N$  where  $Q(m)$  denoted by  $q_m$  is the

probability of moving  $m$  steps to the right of the current location. Throughout this thesis, we use

1. Algorithm “ $Q_y$ ”:

$$q_m := \begin{cases} 1 - 2y, & m = 0 \\ y, & m = 1 \\ y, & m = N - 1 \end{cases}$$

In this case each molecule moves left or right with probability  $y$  or does not move with probability  $1 - 2y$ . See Table: 2.1 (page 19) for standard values used throughout.

When the molecules have the opportunity to decay, there is a chance  $\alpha \in (0, 1)$  that the molecule will leave the system per fast time step. It cannot be observed or consumed by the creature after a decay. See Table: 2.1 (page 19) for standard values.

It will be shown that with repetitions of injection, decay and movement of the molecules but not the injector, the random variable describing the number of molecules at a given location on the ring goes to a steady state. Furthermore, it will be shown that the steady state at a given location is a Poisson random variable defined by its distance from the injector and independent of the number of molecules at each other point on the ring.

### 2.2.5 Creatures

A creature can observe molecules and decide where to move. Let  $C^t$  be the random variable describing the creature’s position on the ring at time  $t$ . Furthermore, let  $M^t$  be the random variable describing the number of molecules at position  $C^t$  at time  $t$ . When a creature makes an observation of the molecules at some time,  $t$ , the creature sees  $M^t$  molecules. Throughout this thesis we will denote the actual location of the creature at time  $t$  and the actual number of molecules at location to be  $c_t$  and  $m_t$  respectively. We will denote  $c_{t+1}$  by the more convenient  $c_+$ .

The sequence of realizations of  $M^t$  and  $C^t$  are the sum total of the creature's interactions with the world. The creature's history up to time  $t$  is the collection of observed molecule counts and creature locations at times  $0 \leq t' \leq t$ . We define a particular history  $h^t$  as that subset of the set of all (infinitely long) sequences of locations and observations that is consistent with a particular finite initial sequence, as follows:

$$h^t := \bigcap_{t'=0}^{t'=t} \left( \{M^{t'} = m_{t'}\} \cap \{C^{t'} = c_{t'}\} \right)$$

or equivalently,

$$h^t = \{M^t = m_t\} \cap \{C^t = c_t\} \cap h^{t-1}.$$

Note that for any  $t > 0$ ,  $h^t \subset h^{t-1}$ . For consistency, for any  $k < 0$  we define  $h^k$  to be the unconditioned set of all histories. We will use the notation  $\mathcal{H}^t$  to denote the random variable corresponding to the creature's history up to and including time  $t$ . Any particular realization of this random variable is some  $h^t$ .

The model creature will be endowed with perfect information about several aspects of its environment. In order to focus on the core issue of the (in)equivalence of information maximization and utility maximization, we will simplify the problem by assuming the creature has complete "knowledge" about every aspect of the world except its actual distance from the injector. Therefore we assume the creature knows the following:

1. The creature knows that there are exactly  $N$  different locations on a 1-dimensional, ring world.
2. The creature knows that at time  $t = 0$  both the injector and itself start at independent, uniformly random locations.
3. The creature knows its history,  $\mathcal{H}^t$ .
4. The creature knows the injector's movement algorithm,  $P$ .
5. The creature knows the dynamics of the molecules.



- (a) The creature knows the injector's injection rate,  $\gamma$ .
- (b) The creature knows the decay rate,  $\alpha$ .
- (c) The creature knows the molecule movement algorithm,  $Q$ .

The last three facts can be replaced with the knowledge of the molecules' steady state distribution about the injector. Although mathematically equivalent in terms of analyzing behaviors it is perhaps easier to imagine evolution sensitizing a species to the distribution of food in the world than the parameters that cause the distribution of food.

Using the creature's inherent knowledge, its observations, and its knowledge of its position when making those observations, the creature can develop a probability mass function (p.m.f.) for the injector's location. This p.m.f. will be a conditional probability distribution; *i.e.* it will be conditioned on the history of the creature's movements and observations. In order to describe the process by which the creature's estimate of the injector location is updated, we introduce the following random variables.

Let  $F_-^t$  be a random variable denoting the conditional probability mass function describing the injector's location at time  $t$ , given the past history  $\mathcal{H}^{t-1}$ , *not including the observation at time  $t$* . Let  $f_-^t$  be the actual conditional probability mass function describing the injector's location at time  $t$ , given the past history  $h^{t-1}$ . The function  $f_-^t(s)$  is the probability that the injector is at position  $s$  at time  $t$ , given the creature's past history. This probability mass function is well defined if and only if  $h^{t-1}$  is in the support of  $\mathcal{H}^{t-1}$  (*i.e.* that particular history occurs with nonzero probability).

Let  $F_+^t$  be a random variable denoting the conditional probability mass function describing the injector's location at time  $t$ , given the full history  $\mathcal{H}^t$ , *including the observation at time  $t$* . Let  $f_+^t$  be the actual conditional probability mass function describing the injector's location at time  $t$ , given the full history  $h^t$ . The function  $f_+^t(s)$  is the probability that the injector is at position  $s$  at time  $t$ , given the creature's

full history  $h^t$ . This probability mass function is well defined if and only if  $h^t$  is in the support of  $\mathcal{H}^t$ . Sections 4.3 and 4.4 detail how  $F_+^t$  is determined by  $F_-^t$  and the molecule count  $M^t$  and the creature's position  $C^t$  at time  $t$ , and how  $F_-^{t+1}$  is determined by  $F_+^t$  and the injector movement algorithm.

In Section 4 it will be shown that unlimited, accurate memory of past observations and locations,  $\mathcal{H}^t$ , can be replaced by accurate memory of  $F_-^t$ ,  $C^t$  and  $M^t$  when calculating  $F_+^t$ , and accurate memory of  $F_+^t$  when calculating  $F_-^{t+1}$ .

At time  $t = 0$  and before any observation, the creature can exploit the knowledge that the injector and itself are uniformly likely to be at any location on the ring and that the injector's location is independent of the creature's. Therefore with probability one,  $F_-^0(l) = 1/N$  for all locations  $l$ . This is equivalent to predicting the injector's location given no information.

The description provided above applies to all types of creature studied in this thesis. What distinguishes one creature from another is its movement strategy: the rule(s) by which it determines how to move in reaction to its observations of the environment. Specifically, the creature's position is determined by  $\mathcal{H}^{t-1}$  and the creature's (possibly stochastic) movement strategy. Section 7 provides examples of specific movement strategies.

### 2.2.6 Event Sequence

We use discrete time steps to update the positions of the injector, the creature and the molecules.

Between any activity in a time step repeating itself (e.g. a creature observing  $M^t$  at time step  $t$  and then observing  $M^{t+1}$  at time step  $t + 1$ ) many other actions take place. It is important that words like current, previous and next be well-defined. Two events occur in the same (slow) time step,  $t$ , if and only if they occur between the injector arriving at some location and being able to move to the next. The following is an ordered list of events that occur in a (slow) time step  $t$ .

1. The injector arrives at its location for this time step.
2. The molecules distribute to steady state.
3. The creature observes the number of molecules,  $M^t$ , at its current location,  $C^t$ .
4. The creature moves and is able to use  $\mathcal{H}^t$  to choose  $C^{t+1}$ .
5. The injector moves according to its movement algorithm,  $P$ .

The molecules distribute under the fast time steps  $\tau$ . The following is an ordered list of events that occur in each (fast) time step  $\tau$ .

1. Each molecule has a chance,  $\alpha$ , to decay and leave the system.
2. Each molecule moves according to the algorithm,  $Q$ .
3. The injector injects  $R$  molecules to its current location.

For technical reasons, time  $\tau$  marks the beginning of this sequence. i.e. At time  $(t : \tau) = (t : 0)$  no molecule decays, movements, or injections have been made during time step  $t$ . In the analytic results, the molecule distribution at the  $\lim \tau \rightarrow \infty$  is considered.

Symbol	Definition	Range	Std. Value
$N$	An even integer representing the size of world	$[18, \infty)$	18
$\gamma$	The mean injection intensity.	$[0, \infty)$	Varied
$P_x^+$	An injector movement algorithm.	$x \in [0, 1/2]$	$(1/4, 1/3)$
$Q_y^{++}$	A molecule movement algorithm.	$y \in (0, 1/2]$	$1/3$
$\alpha$	The probability a molecule decays on iteration of decay.	$(0, 1)$	0.1, 0.4

Table 2.1: Standard Parameter Values. These values describe the world in its entirety excepting only the creature. All other values describing the world are derived from these parameters.  $+$  is fully defined in Section: 2.2.3.  $++$  is fully defined in Section: 2.2.4.

Symbol	Definition
$R^{t:\tau}$	The Poisson random variable that describes the number of molecules the injector delivers on each injection.
$\gamma$	The mean value of $R$ .
$S^t$	The random variable describing the injector's location at time $t$ .
$P, P_x$	Probability mass functions defining a movement algorithm for the injector. Movement is relative to the current position. See Section 2.2.3.
$p_m$	The probability an injector moves $m$ steps to right when it moves.
$Q, Q_y$	Probability mass functions defining a movement algorithm for the molecules. Movement is relative to the current position. See Section 2.2.4.
$q_m$	The probability a molecule moves $m$ steps to right when it moves.
$M^t$	The random variable describing the the number of molecules observed by the creature at time $t$ .
$m_t$	The actual number of molecules observed at time $t$ .
$C^t$	The random variable describing the creature's location at time $t$ .
$c_t$	The creature's actual location at time $t$ .
$c_+$	The creature's actual location at the next time step.
$\mathcal{H}^t$	The random variable describing creature's history of positions and molecules observations up to and including time $t$ .
$h^t$	The creature's actual history of observations and locations up to and including time $t$ .
$K_i^{t:\tau}$	The random variable describing the number of molecules $i$ steps to the right of the injector at time $t : \tau$ . It will be shown that $K_i^{t:\tau}$ is Poisson.
$K_i$	The random variable describing the number of molecules $i$ steps to the right of the injector at steady state. It will be shown that $K_i$ is Poisson.
$\hat{\lambda}(\gamma)$	A vector of mean values of molecule counts, given an injection rate $\gamma$ where $\lambda_i(\gamma) = \mathbb{E}(K_i)$
$\hat{\lambda}$	A vector of means where $\lambda_i = \mathbb{E}(K_i)$ . This symbol is used when the form of an equation or a result is independent of $\gamma$ .

Table 2.2: Random Variables, Probability Mass Functions and Parameters

Symbol	Definition	Page
$\mathbb{E}(arg)$	The expected value of the <i>argument</i> .	N/A
$\mathbb{P}(Event)$	The probability that the <i>Event</i> occurs.	N/A
$:=$	Read as “is defined to be”.	N/A
$\begin{bmatrix} \vdots \end{bmatrix}^*$	The transpose of a vector. $\begin{bmatrix} \vdots \end{bmatrix}^* = [\dots]$ .	N/A
$H(arg)$	The entropy of a random variable or probability mass function <i>arg</i> . Shannon’s discrete entropy measure.	3
$H(arg1 arg2)$	The conditional entropy of one random variable or probability mass function <i>arg1</i> given another <i>arg2</i> .	3
$H(arg1;arg2)$	The mutual information between two random variables or probability mass functions <i>arg1</i> and <i>arg2</i> .	3
$t$	A natural number defining a (slow) time step for the creature and injector. $t \in \mathbb{N}$	6
$\tau$	A natural number defining a (fast) time step for molecules within some time step. $\tau \in \mathbb{N}$	6
$t : \tau$	The beginning of time step $\tau$ within time step $t$ .	6
$N$	The number of locations in the 1-dimensional world.	12
$\mathcal{P}(\mathbb{Z}_N)$	The set of all probability mass functions on $\mathbb{Z}_N$ .	12
$\mathcal{J}_+$	The set of $f_+ \in \mathcal{P}(\mathbb{Z}_N)$ such that there is an $e$ where $f_+(e) = 1$ or a $d$ such that $f_+(e+d) + f_+(e-d) = 1$ . The partitions of $\mathcal{J}_+$ are defined in Section 6.2.	48
$f_+$	A symbol used to denote an arbitrary probability mass function in $\mathcal{J}_+$ .	N/A
$\mathcal{J}_-$	The set of $f_- \in \mathcal{P}(\mathbb{Z}_N)$ that are the support of $F_-^t$ as the injection rate goes to infinity. The partitions of $\mathcal{J}_-$ are defined in Section 6.3.	50
$f_-$	A symbol used to denote an arbitrary probability mass function in $\mathcal{J}_-$ .	N/A

Table 2.3: Variables and Notation, their definition and the page on which they are introduced.

## 3 Molecules in the World

### 3.1 Overview

This section derives many of the properties of the steady state distribution of the molecules. The steady state is reached by iterations of decay, molecule movement and injection while the injector is not moving as described in Section: 2.2.6. The following is a list of important results proved in this section:

1. The most important of these results is that there is in fact a steady state. In a stochastic system steady state means that the probability of observing  $k$  molecules at location  $i$  does not change over time.
2. The number of molecules at any location and any time is a Poisson random variable.
3. Because decay and molecule movement  $Q$ , are independent of the molecules' location, but injection is dependent upon the injector's location, the number of molecules at a given location is a function of the decay rate,  $Q$ , the injection rate, and the injector's location.
4. For the molecule movement algorithm  $Q_y$ , the average number of molecules at the injector's location is greater than the average number of molecules one step away from injector's location, and decreases monotonically with increasing distance from the injector.

The molecules, injector, and creature reside on an  $N$  unit, one-dimensional circular world. On each (fast) time step,  $\tau$ , each molecule moves according to the movement algorithm  $Q$ . In addition each molecule has a probability  $\alpha$  of vanishing due to an unspecified decay process. The injector then injects  $R^{t:\tau}$  molecules at the injector's position, where  $R^{t:\tau}$  is a Poisson random variable with mean  $\gamma$ . See Table: 2.1 for the standard parameter values. The number of molecules at any given location at

any time step  $(t : \tau)$  can be described by a random variable that is based on these parameters. In order keep the the following work as general as possible and avoid being bogged down in notation the actual parameters affecting the quantitative but not qualitative nature of the random variables will not be listed each time the random variables are mentioned.

Let  $K_i^{t:\tau}$  be the random variable describing the number of molecules  $i$  steps to the right of the injector at time  $(t : \tau)$ .

Let  $K_i^t$  be the  $\lim_{\tau \rightarrow \infty} K_i^{t:\tau}$ . It will be shown that  $K_i^t$  is a Poisson random variable with a mean  $\lambda_i$ . Furthermore, it will be shown that  $\lambda_i$  is independent of  $t$  and therefore  $K_i^t$  will simply be referred to as  $K_i$

Before any injections are made into the world, every position has 0 molecules. It is equivalent to say that the number of molecules at any location is a Poisson random variable with mean 0.

$$\mathbb{P}(K_i^{0:0} = k) = \frac{e^0(0)^k}{k!} = \begin{cases} 1 & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

One of the major goals of this section is to show that there is a steady state distribution of molecules. We have shown  $K_i^{0:0}$  is a Poisson random variable. In Sections: 3.2 and 3.3, we will show the processes of decay, molecule movement and injection maintain the Poisson nature of  $K_i^{0:\tau}$ . Induction then implies  $K_i^{0:\tau}$  is a Poisson random variable  $\forall \tau \in \mathbb{N}$ . Furthermore, in Section: 3.4 we will show the  $\lim_{\tau \rightarrow \infty} K_i^{0:\tau}$  is a Poisson random variable. From there, we will be able to prove  $\lim_{\tau \rightarrow \infty} K_i^{t:\tau} \forall t$  is a Poisson random variable.

### 3.2 Decay/Movement with Poisson Random Variables

Let  $X$  be a Poisson random variable with mean  $\lambda_x$  representing the number of molecules at a given location, and let  $X_1$  be the number of molecules remaining after an iteration of decay. Each molecule can decay with a probability of  $0 < \alpha < 1$ .

**Lemma 3.1.**  $X_1$  is a Poisson random variable with mean  $(1 - \alpha)\lambda_x$ .

*Proof.*  $\mathbb{P}(X_1 = x_1)$  is the probability that there were  $x_1$  molecules,  $\mathbb{P}(X = x_1)$ , and that no molecules decayed plus the probability that there were  $x_1 + 1$  molecules,  $\mathbb{P}(X = x_1 + 1)$ , and that one molecule decayed and so on.

$$\begin{aligned}
\mathbb{P}(X_1 = x_1) &= \sum_{j=0}^{\infty} \mathbb{P}(X = x_1 + j) \binom{x_1 + j}{j} (1 - \alpha)^{x_1} \alpha^j \\
&= \sum_{j=0}^{\infty} \frac{\mathbb{P}(X = x_1) \lambda_x^j x_1!}{(x_1 + j)!} \binom{x_1 + j}{j} (1 - \alpha)^{x_1} \alpha^j \\
&= \frac{e^{-\lambda_x} ((1 - \alpha)\lambda_x)^{x_1}}{x_1!} \sum_{j=0}^{\infty} \frac{(\alpha\lambda_x)^j}{j!} \\
&= \frac{e^{-((1-\alpha)\lambda_x)} ((1 - \alpha)\lambda_x)^{x_1}}{x_1!}
\end{aligned}$$

The number of molecules remaining after each molecule can move away with probability  $\alpha$  is also a Poisson random variable with mean  $(1 - \alpha)\lambda$ . The proof for this is identical to the proof for the decay.  $\square$

**Lemma 3.2.** The number of molecules moved,  $J$ , with probability  $q \in [0, 1]$  from  $X$  is a Poisson random variable with mean  $q\lambda_x$ .

*Proof.*  $\mathbb{P}(J = j)$  is the probability that there were  $j$  molecules,  $\mathbb{P}(X = j)$ , and that  $j$  molecules move plus the probability that there were  $j + 1$  molecules,  $\mathbb{P}(X = j + 1)$ , and that  $j$  molecules move and so on.

$$\begin{aligned}
\mathbb{P}(J = j) &= \sum_{x=j}^{\infty} \mathbb{P}(X = x) \binom{x}{j} (1 - q)^{x-j} q^j \\
&= \sum_{y=0}^{\infty} \mathbb{P}(X = y + j) \binom{y + j}{j} (1 - q)^y q^j && (y = x - j) \\
&= \frac{e^{-\lambda_x} (q\lambda_x)^j}{j!} \sum_{y=0}^{\infty} \frac{((1 - q)\lambda_x)^y}{y!} \\
&= \frac{e^{-q\lambda_x} (q\lambda_x)^j}{j!} && \square
\end{aligned}$$



We have now shown that when molecules decay at or move away from some location  $i$ , the Poisson nature of  $K_i^{0:\tau}$  is maintained. We have also shown the number of molecules moved by the molecule movement algorithm is a Poisson random variable.

### 3.3 The Sum of Two Poisson Random Variables

To show that movement and injection maintain the Poisson nature of the molecule counts we need to show that the addition of a Poisson number of molecules maintains the Poisson nature of the molecule counts.

Let  $X$  and  $Y$  be independent Poisson random variables with means  $\lambda_x$  and  $\lambda_y$  respectively.

**Lemma 3.3.**  $Z = X + Y$  is a Poisson random variable with mean  $(\lambda_x + \lambda_y)$ .

*Proof.*

$$\begin{aligned}
\mathbb{P}(Z = z) &= \mathbb{P}(X + Y = z) \\
&= \sum_{j=0}^z \mathbb{P}(X = j \cap Y = z - j) \\
&= \sum_{j=0}^z \mathbb{P}(X = j)P(Y = z - j) && \text{(using independence)} \\
&= \sum_{j=0}^z \frac{e^{-\lambda_x} \lambda_x^j}{j!} \frac{e^{-\lambda_y} \lambda_y^{z-j}}{(z-j)!} \\
&= \frac{e^{-(\lambda_x + \lambda_y)}}{z!} \sum_{j=0}^z \binom{z}{j} \lambda_x^j \lambda_y^{z-j} \\
&= \frac{e^{-(\lambda_x + \lambda_y)} (\lambda_x + \lambda_y)^z}{z!} && \text{(using binomial formula)} \quad \square
\end{aligned}$$

The above proves that Poisson injections and molecules movements maintain the Poisson nature of the molecule count at any location.

### 3.4 Molecule Counts are Poisson

**Theorem 3.4.**  $K_i^0$  is a Poisson random variable.

*Proof.* The molecules injected  $n$  (fast) iterations ago have undergone one injection,  $n$  decays and  $n$  moves. For the rest of this section a time step or iteration may be assumed to refer to the fast molecule time steps denoted by  $\tau$  unless specifically noted otherwise. As shown above (Lemmas 3.1, 3.2 and 3.3), the processes of injection, decay, and movement maintain the Poisson nature of the molecule counts at each location.

According to [9, Countable Additivity Theorem], if the sum of the means of Poisson random variables converges to  $\sigma$ , then the sum of the random variables is a Poisson random variable with mean  $\sigma$ .

Let  $\mu_{in}$  be the mean for the Poisson random variable that expresses the number of molecules at position  $i$  that were injected  $n$  time steps ago.

Let  $\lambda_{in} = \sum_{j=0}^n \mu_{ij}$  be the mean describing the (Poisson) number of molecules at position  $i$  that were injected within the last  $n$  time steps.  $\mu_{in} \geq 0 \forall i, n$  and  $\forall n \in \mathbb{N}$  which implies that  $\lambda_{i(n+1)} \geq \lambda_{in} \geq 0 \forall i \forall n \in \mathbb{N}$ . If there is an upper bound for  $\lim_{n \rightarrow \infty} \lambda_{in} = \lambda_i$ , then the limit converges by the Monotone Convergence Theorem.

The mean describing the (Poisson) number of surviving molecules in the entire world that were injected  $n$  time steps ago is simply  $(1 - \alpha)^n \gamma$  where  $1 > \alpha > 0$  is the decay rate and  $\gamma$  is the mean of the molecules injected. For a given  $n$ ,  $\sum_i \mu_{in}$  is the mean describing the same random variable. This implies that  $\mu_{in} \leq (1 - \alpha)^n \gamma \Rightarrow \lambda_{in} \leq \sum_{j=0}^n (1 - \alpha)^j \gamma < \frac{\gamma}{\alpha} < \infty, \forall n \in \mathbb{N}$ . So  $\lambda_{in} \rightarrow \lambda_i \leq \frac{\gamma}{\alpha}$ , as  $n \rightarrow \infty$ .

Therefore,  $K_i^0$ , is a Poisson random variable with mean  $\lambda_i$ . □

This proves that the molecules injected in any (slow) time step  $t$  form Poisson distributions at each location  $i$ . It still remains to show what happens to the molecules that were injected in (slow) time steps  $t' \in \{1, \dots, t - 1\}$ .

**Lemma 3.5.** *The number of molecules in the system at any time ( $t : \tau$ ) is finite.*

*Proof.* This is true because at time  $(0 : 0)$  there are finitely many  $(0)$  molecules in the world.

After finitely many injections,  $n$ , the number of molecules in the world is a Poisson random variable with mean  $\sigma_n = \sum_{j=0}^n (1 - \alpha)^j \gamma$ .

After infinitely many injections,  $\lim_{n \rightarrow \infty} \sigma_n = \gamma / \alpha$ . Using [9, Countable Additivity Theorem], the total number of molecules is Poisson a random variable with mean  $\gamma / \alpha$ .  $\square$

**Corollary 3.6.**  $K_i^t$  is a Poisson random variable with mean  $\lambda_i$  and is independent of  $t$ .

*Proof.* From above, there exists a  $\lambda_i$  for each location  $i$ , that is the Poisson parameter of  $K_i^0$ .  $K_i^0$  describes the molecule counts when starting with no molecules in the system. So,  $K_i^t$  is  $K_i^0$  plus whatever molecules were present at time  $(t : 0)$  that are still present and are  $i$  steps to the right of the injector.

Suppose there is a molecule in the system at time  $(t : 0)$ . The probability that the molecule has not decayed at time  $(t : \tau)$  is  $(1 - \alpha)^\tau \gamma$ . The  $\lim_{\tau \rightarrow \infty} (1 - \alpha)^\tau \gamma = 0$ . This means that a molecule in the world at time  $(t : 0)$  has no probability of surviving in the limit. There are finitely many molecules at time  $(t : 0)$  and so the probability that any of the molecules in existence at time  $(t : 0)$  survive in the limit is also 0. Therefore the distribution of molecules at the beginning of a slow time step  $(t : 0)$  has no effect on the distribution of molecules at the end of that slow time step  $(t : \tau \rightarrow \infty)$ .

Consequently  $K_i^t$  and  $K_i^0$  are Poisson random variables with the same mean,  $\lambda_i$ .  $\square$

### 3.5 Calculating $\mathbb{E}(K_i)$ ( $\lambda_i$ )

Although it has been shown that the number of molecules  $i$  steps to the right of the injector at time  $t$ ,  $K_i$ , is a Poisson random variable with a mean  $\lambda_i$ , a method for calculating  $\lambda_i$  has not yet been presented. As shown in Section: 3.4,  $\lambda_{in}$  is the mean of the Poisson distribution of molecules  $i$  steps to the right of injector that were injected within the last  $n$  (fast) time steps. For the rest of this section a time step or

iteration may be assumed to refer to the fast molecule time steps denoted by  $\tau$  unless specifically noted otherwise.  $\lambda_{in}$  is the mean of the Poisson random variable  $K_i^{0:n}$  because the only molecules in the system at time 0 :  $n$  are those that were injected within the last  $n$  time steps.

Let  $\hat{\lambda}_n$  be the  $N \times 1$  vector of  $\lambda_{in}$  for each  $i \in \{0, \dots, N-1\}$ .

Let  $\hat{\lambda}$  be the  $N \times 1$  vector of  $\lambda_i$  for each  $i \in \{0, \dots, N-1\}$ .

Let  $\hat{\gamma}$  be the  $N \times 1$  vector with  $\gamma$  as the first entry and the rest 0's.

$$\hat{\lambda}_n := \begin{bmatrix} \lambda_{0n} \\ \lambda_{1n} \\ \vdots \\ \lambda_{(N-1)n} \end{bmatrix}, \hat{\lambda} := \begin{bmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{(N-1)} \end{bmatrix}, \hat{\gamma} := \begin{bmatrix} \gamma \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Let

$$\tilde{Q} := \begin{bmatrix} q_0 & q_{N-1} & \cdots & q_2 & q_1 \\ q_1 & q_0 & \cdots & q_3 & q_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ q_{N-2} & q_{N-3} & \cdots & q_0 & q_{N-1} \\ q_{N-1} & q_{N-2} & \cdots & q_1 & q_0 \end{bmatrix}.$$

Let  $\tilde{Q}_{lm}$  with  $l, m \in \{1, \dots, N\}$  be the  $l^{\text{th}}$  element of the  $m^{\text{th}}$  column of  $\tilde{Q}$ .  $\tilde{Q}_{lm}$  is the probability a molecule  $m-1$  steps to the right of the injector moves to the location  $l-1$  steps to the right of the injector. Each entry in the matrix comes from  $Q$ , the molecule movement algorithm. The symmetry is due to the fact that the molecule movement algorithm is independent of the molecules starting position.

In Sections 3.2 and 3.3 the effect of decay, molecule movement, and injection on  $\hat{\lambda}_n$  is laid out. After an iteration of decay with probability  $\alpha$ ,  $\hat{\lambda}_n$  becomes  $(1-\alpha)\hat{\lambda}_n$ . Then after the molecules are able to move with a movement algorithm  $Q$ ,  $(1-\alpha)\hat{\lambda}_n$  becomes  $(1-\alpha)\tilde{Q}\hat{\lambda}_n$ . Finally an injection of  $R$  molecules with mean  $\gamma$  implies  $(1-\alpha)\tilde{Q}\hat{\lambda}_n$  becomes  $(1-\alpha)\tilde{Q}\hat{\lambda}_n + \hat{\gamma} = \hat{\lambda}_{n+1}$ . Because  $\hat{\lambda} = \lim_{n \rightarrow \infty} \hat{\lambda}_{n+1}$ ,  $(1-\alpha)\tilde{Q}\hat{\lambda} + \hat{\gamma} = \hat{\lambda}$  or

equivalently,

$$((1 - \alpha)\tilde{Q} - I)\hat{\lambda} = -\hat{\gamma} \quad (3.5.1)$$

where  $I$  is the  $N \times N$  identity matrix.

Let  $v \in \mathbb{R}^N$  be a solution to Equation (3.5.1). Section 3.4 demonstrates that  $((1 - \alpha)\tilde{Q} - I)v = -\hat{\gamma}$  has a solution ( $v = \hat{\lambda}$ ). It remains to be seen if the solution is unique.

**Proposition 3.7.**  $(1 - \alpha)\tilde{Q} - I$  is non-singular and hence  $\hat{\lambda}$  is a unique solution.

*Proof.*  $\sum_m \tilde{Q}_{lm} = 1$  because every molecule that was just moved must have come from some location with a probability of 1.  $\tilde{Q}_{lm} > 0$  because  $\tilde{Q}_{lm}$  is a probability. And  $1 > 1 - \alpha > 0$  because  $1 > \alpha > 0$ . Therefore,  $1 > (1 - \alpha)\sum_m \tilde{Q}_{lm} > 0$ .

Subtracting  $(1 - \alpha)\tilde{Q}_{ll}$  from both sides of  $1 > (1 - \alpha)\sum_m \tilde{Q}_{lm}$  and taking the absolute value results in  $|(1 - \alpha)\tilde{Q}_{ll} - 1| > (1 - \alpha)|\sum_{m:l \neq m} \tilde{Q}_{lm}|$ . With the triangle inequality  $|(1 - \alpha)\tilde{Q}_{ll} - 1| > (1 - \alpha)\sum_{m:l \neq m} |\tilde{Q}_{lm}|$ .

Each diagonal entry of the matrix  $((1 - \alpha)\tilde{Q} - I)$  is  $(1 - \alpha)\tilde{Q}_{ll} - 1$  and the sum of all entries of any row excluding the diagonal is  $(1 - \alpha)\sum_{m:l \neq m} |\tilde{Q}_{lm}|$ . Therefore  $(1 - \alpha)\tilde{Q} - I$  is a diagonally dominant matrix. By the Levy-Desplanques theorem [14],  $(1 - \alpha)\tilde{Q} - I$  is non-singular.  $\square$

Because  $(1 - \alpha)\tilde{Q} - I$  is non-singular, it is invertible. Therefore,

$$\hat{\lambda} = -((1 - \alpha)\tilde{Q} - I)^{-1}\hat{\gamma} \quad (3.5.2)$$

### 3.6 Properties of the Mean Concentration Profile $\{\lambda_i\}_{i=0}^{N-1}$

In this section we will show, for the standard molecule movement algorithm  $Q_y$ ,

1.  $i = 0$  is a local maximum of the mean equilibrium molecule count vector  $\hat{\lambda} = (\lambda_{1-N/2}, \dots, \lambda_{-1}, \lambda_0, \lambda_1, \dots, \lambda_{N/2})$ ,
2.  $\forall i : 0 \leq i < N/2, \lambda_i > \lambda_{i+1}$ ,

3.  $\forall i : 0 \leq i < N/2, \lambda_{-i} > \lambda_{-(i+1)},$
4.  $\forall i : 0 \leq i < N/2 - 1, \lambda_i - \lambda_{i+1} > \lambda_{i+1} - \lambda_{i+2},$
5.  $i = 0$  is a unique local maximum of  $\hat{\lambda}$ , and
6.  $\lambda_j = \lambda_i \iff i = \pm j,$

from which we conclude that the closer a location is to the injector, the greater the expected amount of food.

These results are not required for Section 4 but will play a role in the analytic results to be presented in Section 5.

So far this thesis has dealt with a very general  $Q$ . Let us consider the case where  $Q$  is symmetric. A molecule movement algorithm is defined to be *symmetric* when  $\forall i$   $q_i = q_{-i}$ . This is to say that the probability of moving  $i$  steps to the right is the same as moving  $i$  steps to the left. One can easily see that  $Q_y$  is symmetric (See Section: 2.2.4).

**Lemma 3.8.** *If  $Q$  is symmetric then  $\hat{\lambda}$  is symmetric (i.e.  $\forall i \lambda_i = \lambda_{-i}$ ).*

*Proof.* Every molecule that could be observed by the creature in the slow time step  $t$  was injected at the source's location during time step  $t$  (See Section: 3.4). Any molecule at location  $i$  when the creature makes its observation has survived some number,  $n$ , of decays and moved  $n$  times.

Because  $q_i = q_{-i}$  it is equally probable that the molecule could have survived  $n$  iterations of decay and moved the same distance but in the opposite direction for each move. Consequently it is equally probable the molecule could be at location  $-i$ . This is true of any molecule at any location injected at any time. Therefore,  $\hat{\lambda}$  is symmetric. □

A molecule movement algorithm,  $Q$ , is defined to be *connected* when it is possible for a molecule at any location on the ring to move to any other location.  $Q_y$  is

connected because a molecule that moves one step to the right  $N$  times visits each location on the ring.

**Lemma 3.9.** *If  $Q$  is connected,  $\forall i \lambda_i > 0$ .*

*Proof.* Assume without loss of generality that the injector is at location 0. If  $Q$  is connected, it is possible for a molecule at any location on the ring to move to any other location. Therefore, it is possible after  $N$  fast time steps a molecule could be at any location  $i$ . This implies the probability of observing 1 molecule at location  $i$ ,  $\mathbb{P}(K_i = 1)$ , is greater than zero.  $K_i$  is Poisson (See Corollary 3.6) so  $\mathbb{P}(K_i < 0) = 0$ . Finally,  $\lambda_i = \mathbb{E}(K_i) = \sum_k \mathbb{P}(K_i = k)k > \mathbb{P}(K_i = 1) > 0$ .  $\square$

A local maximum of  $\hat{\lambda}$  is a location  $i^*$  where  $\lambda_{i^*-1} \leq \lambda_{i^*} \geq \lambda_{i^*+1}$ .

**Lemma 3.10.** *For  $Q_y$  with  $0 < y \leq 1/2$ ,*

1. *the injector's location ( $i = 0$ ) is a unique local maximum of  $\hat{\lambda}$ ,*
2.  $\lambda_{|i|} > \lambda_{|i+1|} \forall i < N/2$
3.  $\lambda_j = \lambda_i \iff i = \pm j$

*Proof.* First, we will show that the injector's location ( $i = 0$ ) is a local maximum of  $\hat{\lambda}$ . Turning our attention to Equation: (3.5.1),

$$\begin{aligned}
-\gamma &= (1 - \alpha)(\lambda_0 q_0 + \lambda_{N-1} q_1 + \lambda_1 q_{N-1}) - \lambda_0 \\
&= (1 - \alpha)(q_0 \lambda_0 + 2\lambda_1 q_1) - \lambda_0 \\
&= (1 - \alpha)((1 - 2y)\lambda_0 + 2y\lambda_1) - \lambda_0 \\
&= 2y(1 - \alpha)(\lambda_1 - \lambda_0) + (1 - \alpha)\lambda_0 - \lambda_0 \\
&= 2y(1 - \alpha)(\lambda_1 - \lambda_0) - \alpha\lambda_0
\end{aligned} \tag{3.6.1}$$

The second equality holds because  $Q_y$  is symmetric which implies  $\hat{\lambda}$  is symmetric. The third equality holds by definition of  $Q_y$ .

Moving  $\alpha\lambda_0$  to the other side of the equality in Equation (3.6.1) implies

$$\begin{aligned}
2y(1-\alpha)(\lambda_1 - \lambda_0) &= \alpha\lambda_0 - \gamma \\
&= \alpha\lambda_0 - \alpha\frac{\gamma}{\alpha} \\
&= \alpha(\lambda_0 - \sum_j \lambda_j) \\
&= -\alpha \sum_{j=1}^{j=N-1} \lambda_j \\
&< 0
\end{aligned}$$

The third equality holds because  $\sum_i \lambda_i = \gamma/\alpha$  as shown in Section 3.4. The final inequality holds because  $Q_y$  is connected and so  $\forall i \lambda_i > 0$ .

$2y(1-\alpha) > 0$ , so  $\lambda_1 - \lambda_0 < 0$ . Specifically,

$$\lambda_1 - \lambda_0 = \frac{-\alpha}{2y(1-\alpha)} \sum_{j=1}^{j=N-1} \lambda_j \quad (3.6.2)$$

Therefore,  $\lambda_{N-1} \leq \lambda_0 \geq \lambda_1$  and the injector's location is a local maximum.

Having shown  $\lambda_0 > \lambda_1$  we will use induction to show  $\lambda_i > \lambda_{i+1} \forall i < N/2$ . The Inductive Hypothesis is  $\forall i \ 1 \leq i \leq N/2$

$$\lambda_i - \lambda_{i-1} = \frac{-\alpha}{2y(1-\alpha)} \sum_{j=i}^{j=N-i} \lambda_j \quad (3.6.3)$$

Turning our attention to Equation: (3.5.1) for all  $i \neq 0$ ,

$$\begin{aligned}
0 &= (1-\alpha)(\lambda_i q_0 + \lambda_{i-1} q_1 + \lambda_{i+1} q_{N-1}) - \lambda_i \\
&= (1-\alpha)((1-2y)\lambda_i + y\lambda_{i-1} + y\lambda_{i+1}) - \lambda_i \\
&= y(1-\alpha)(\lambda_{i+1} - \lambda_i) + (1-\alpha)((1-y)\lambda_i + y\lambda_{i-1}) - \lambda_i
\end{aligned}$$

Dividing by  $y(1-\alpha)$  and moving the term  $(\lambda_{i+1} - \lambda_i)$  to the other side of the equality



results in

$$\begin{aligned}
(\lambda_{i+1} - \lambda_i) &= \frac{-(1-\alpha)((1-y)\lambda_i + y\lambda_{i-1}) + \lambda_i}{y(1-\alpha)} \\
&= \frac{\alpha\lambda_i + y(1-\alpha)(\lambda_i - \lambda_{i-1})}{y(1-\alpha)} \\
&= \frac{\alpha\lambda_i - \frac{\alpha}{2} \sum_{j=i}^{j=N-i} \lambda_j}{y(1-\alpha)} && \text{(Inductive Hypothesis)} \\
&= \frac{\frac{\alpha}{2}(\lambda_i + \lambda_{-i}) - \frac{\alpha}{2} \sum_{j=i}^{j=N-i} \lambda_j}{y(1-\alpha)} && \lambda_i = \lambda_{-i} \text{ (Symmetry)} \\
&= \frac{-\alpha}{2y(1-\alpha)} \sum_{j=i+1}^{j=N-(i+1)} \lambda_j
\end{aligned}$$

This completes the inductive proof. This proof does not imply  $\lambda_{N/2} > \lambda_{N/2+1}$  because  $\lambda_{N/2+1}$  and  $\lambda_{N/2-1}$  have already been subtracted from the  $\sum \lambda_j$ . It does however imply  $\lambda_{N/2-1} > \lambda_{N/2}$ . Therefore  $\lambda_{N/2}$  is *not* a local maximum of  $\hat{\lambda}$ .

We have now shown for  $Q_y$  that

1.  $i = 0$  is a local maximum of  $\hat{\lambda}$
2.  $\lambda_i > \lambda_{i+1} \forall i < N/2$
3.  $\lambda_i - \lambda_{i+1} > \lambda_{i+1} - \lambda_{i+2} \forall i < N/2 - 1$

Due to the symmetry of  $Q_y$  this also means  $\lambda_{-i} > \lambda_{-(i+1)} \forall i < N/2$ . Therefore  $i = 0$  is a unique local maximum of  $\hat{\lambda}$ ,  $\lambda_j = \lambda_i \iff i = \pm j$ , and the closer a location is to the injector the greater the expected amount of food. Finally, the difference of the means of two adjacent locations is smaller the further away from the injector the locations are. □

In Lemma 3.10 the difference between  $\lambda_i$  and  $\lambda_{i+1}$  was calculated. One of properties of that difference is as follows.

**Corollary 3.11.**  $\forall i < N/2 - 1 \lambda_{|i|} - \lambda_{|i+1|} > \lambda_{|i+1|} - \lambda_{|i+2|}$

*Proof.* In Lemma 3.10 it was shown that for all  $i$ ,  $0 \leq i < N/2 - 1$ ,

$$\lambda_i - \lambda_{i+1} > \lambda_{i+1} - \lambda_{i+2}.$$

Using the symmetry of  $Q_y$ ,  $\forall i < N/2 - 1 \lambda_{-i} - \lambda_{-i-1} > \lambda_{-i-1} - \lambda_{-i-2}$ . □

### 3.7 The Effect on $\{K_i\}_{i=0}^{N-1}$ as $\gamma \rightarrow \infty$

Above we studied the effects of  $Q_y$  on  $\hat{\lambda}$ . Here we study the effects as  $\gamma \rightarrow \infty$  on the concentration profile. This section will show that as  $\gamma \rightarrow \infty$  the probability that a given number of molecules observed at location  $i$  could also have been observed at  $j$  goes to 0 when  $\lambda_i \neq \lambda_j$ . The first step in the argument is to show how  $\hat{\lambda}$  is affected by  $\gamma$ .

Let  $\hat{\lambda}(\gamma)$  denote the vector of mean concentrations when the average injection rate is  $\gamma$ .

**Lemma 3.12.** *The vector  $\hat{\lambda}(\gamma) = \gamma\hat{\lambda}(1)$ .*

*Proof.* Let  $\hat{v} \in \mathbb{R}^N$  be the  $N \times 1$  vector with the first entry being 1 and all others being zero. Equation: (3.5.1) shows

$$\begin{aligned} -\hat{\gamma} &= ((1 - \alpha)\tilde{Q} - I)\hat{\lambda}(\gamma) \\ -\gamma\hat{v} &= ((1 - \alpha)\tilde{Q} - I)\hat{\lambda}(\gamma) \\ -\hat{v} &= ((1 - \alpha)\tilde{Q} - I)\left(\gamma^{-1}\hat{\lambda}(\gamma)\right) \end{aligned}$$

$\hat{v}$  is  $\hat{\gamma}$  when  $\gamma = 1$  and  $(1 - \alpha)\tilde{Q} - I$  is invertible so,

$$\gamma^{-1}\hat{\lambda}(\gamma) = \hat{\lambda}(1)$$

$$\hat{\lambda}(\gamma) = \gamma\hat{\lambda}(1) \quad \square$$

Having shown that  $\hat{\lambda}(\gamma)$  is simply  $\gamma\hat{\lambda}(1)$ , it is clear that changing the injection rate does not change the normalized distribution of Poisson random variable means about the injector. However, it does have a significant impact on how much an observation can tell the creature about the injector's location. As can be seen in Figure 3.1, the increasing the injection rate increases the number of standard deviations each mean is from each other mean. This decreases the uncertainty a creature has about the location of the injector from observing a certain number of molecules. For example, in Figure 3.1, with an injection rate of 1 there are seven different locations where

observing 1 molecule is within one standard deviation of the mean, but with an injection rate of 27 there are only two locations where observing 27 molecules is within one standard deviation of the mean. This argument is presented only to provide some insight into the problem.

As  $\gamma \rightarrow \infty$ , the standard deviation of  $K_i$  divided by the injection rate goes to 0.

**Corollary 3.13.** *The standard deviation of  $K_i$  is  $\sqrt{\gamma\lambda_i(1)}$ .*

*Proof.* The standard deviation of a Poisson random variable with mean  $\lambda_i(\gamma)$  is  $\sqrt{\lambda_i(\gamma)}$ . From Lemma 3.12  $\sqrt{\lambda_i(\gamma)} = \sqrt{\gamma\lambda_i(1)}$ . □

Figure 3.1: A plot of the mean number of particles at each location relative to the injector, divided by the injection rate using our standard values ( $\alpha = .1$ ). The error bars represent one standard deviation about the mean, also divided by the injection rate. As  $\gamma$  goes to infinity the number of standard deviations between adjacent means is unbounded.

## 4 Using Sense Data

### 4.1 Overview

This section provides the mathematics to convert histories into the probabilities describing the injector’s next location. The essential idea is that the creature starts with some initial probability mass function describing the injector’s location and then uses subsequent observations and knowledge of the molecules’ distribution and the injector’s movement algorithm to create a p.m.f to describe the injector’s next location.

An interesting consequence of these calculations is that although we assumed that the creature “knows”

- its entire history,  $\mathcal{H}^t$ ,
- the rate of molecule injection,  $\gamma$ ,
- the rate of decay,  $\alpha$ ,
- the molecule movement algorithm,  $Q$ , and
- the injector movement algorithm,  $P$ ,

it is *sufficient* to know only

- the latest observation,  $M^t$  at  $C^t$ ,
- the steady state distribution of molecules,  $\forall i K_i$  is Poisson and  $\mathbb{E}(K_i) = \lambda_i$ ,
- the injector’s movement algorithm,  $P$ , and
- the p.m.f. describing the injector’s location before the latest observation,  $f_-^t$ .

The second set of assumptions is far weaker in that it doesn’t require an infinite memory nor inherent knowledge of molecule dynamics that may be impossible for a creature to discern or inherit through observation or evolution.

In our model, the number of molecules  $i$  steps to the right of injector,  $K_i$ , is Poisson with mean  $\lambda_i$ . This section is just as valid for a model where  $K_i$  is a random variable of some other description. So long as the parameters of each  $K_i$  are known and independent of the creature's history,  $\mathcal{H}^t$ , the results of this section are still applicable.

## 4.2 Review of Notation

As detailed in Sections 2.2.3, 2.2.4, 2.2.5 and 3, we adopt the following notation:

$C^t$  : The random variable describing the creature's position at time  $t$ .

$M^t$  : The random variable describing the number of molecules the creature observes at time  $t$ .

$K_i$  : The number of molecules  $i$  steps to the right of the injector.

$\lambda_i$  : The mean value of the Poisson random variable  $K_i$ .

$S^t$  : The random variable describing the injector's location at time  $t$ .

$\mathcal{H}^t$  : The random variable describing creature's observations and locations up to and including time  $t$ .

$f_-^t$  : The probability mass function for the injector's location at time  $t$  conditioned on  $\mathcal{H}^{t-1}$  and the system parameters.

$f_+^t$  : The probability mass function for the injector's location at time  $t$  conditioned on  $\mathcal{H}^t$  and the system parameters.

In addition we introduce  $\mathcal{M}$  to denote the event that the random variable  $M^t$  is  $m_t$ ,  $\mathcal{C}$  to denote the event that  $C^t$  is  $c_t$  and  $\mathbf{H}^t$  to denote the event  $\mathcal{H}^t$  is  $h^t$ . The set of all events consistent with a history up to time  $t$  is a subset of the set of all events consistent with a history up to time  $(t - 1)$ . Adopting set theoretic notation, we pass

from  $\mathbf{H}^{t-1}$  to  $\mathbf{H}^t$  by taking the *intersection* of the following subsets of all possible events:

$$\mathbf{H}^t = \mathbf{H}^{t-1} \cap \mathcal{M} \cap \mathcal{C} \quad (4.2.1)$$

For consistency of notation we consider  $\mathcal{H}^{-1}$  to be the set of all possible histories (not yet conditioned on any observations).

Before accounting for the first observation, the creature knows the injector is at a random location on the ring. Therefore  $\forall l f_-^0(l) = 1/N$  (See Proposition 2.1).

### 4.3 The Injector's Current Location, Given a Full History

The p.m.f.  $f_+^t$  is defined as  $f_+^t(s) = P(S^t = s | \mathbf{H}^t)$  for  $0 \leq s \leq N - 1$ . For the given  $s$ , let  $\mathcal{S}$  denote the event  $S^t = s$ .

$$\begin{aligned} f_+^t(s) &= \mathbb{P}(\mathcal{S} | \mathbf{H}^t) \\ &= \frac{\mathbb{P}(\mathcal{S} \cap \mathcal{M} \cap \mathcal{C} \cap \mathbf{H}^{t-1})}{\mathbb{P}(\mathcal{M} \cap \mathcal{C} \cap \mathbf{H}^{t-1})} \end{aligned} \quad (4.3.1)$$

$$= \frac{\mathbb{P}(\mathcal{M} | \mathcal{C} \cap \mathcal{S} \cap \mathbf{H}^{t-1}) \cdot \mathbb{P}(\mathcal{S} | \mathcal{C} \cap \mathbf{H}^{t-1}) \cdot \mathbb{P}(\mathcal{C} \cap \mathbf{H}^{t-1})}{\mathbb{P}(\mathcal{M} | \mathcal{C} \cap \mathbf{H}^{t-1}) \cdot \mathbb{P}(\mathcal{C} \cap \mathbf{H}^{t-1})} \quad (4.3.2)$$

Equations (4.3.1) and (4.3.2) are obtained by definition of conditional probability:  $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$ . Note that in Equation 4.3.1,  $\mathbb{P}(B) = \mathbb{P}(\mathbf{H}^t) > 0$  since  $h^t$  refers to a history in the support of  $\mathcal{H}^t$ .

Because the number of molecules observed at the location  $c$  is dependent on the location of the injector but independent of the number of molecules in the system at earlier times (See Corollary 3.6), we have

$$\mathbb{P}(\mathcal{M} | \mathcal{C} \cap \mathcal{S} \cap \mathbf{H}^{t-1}) = \mathbb{P}(\mathcal{M} | \mathcal{C} \cap \mathcal{S}).$$

The probability of observing  $m_t$  molecules given that the creature is at  $c$  and the injector is at location  $s$  is simply  $\mathbb{P}(K_{c-s} = m_t)$ . This is because  $(c - s \bmod N)$  is

the number of steps  $c$  is to the right of  $s$ .

$$\mathbb{P}(\mathcal{M}|\mathcal{C} \cap \mathcal{S}) = \mathbb{P}(K_{c-s} = m_t) \quad (4.3.3)$$

$$= \frac{e^{-\lambda_{c-s}} \lambda_{c-s}^{m_t}}{m_t!} \quad (4.3.4)$$

The probability of the intersection of  $\mathcal{C}$  and  $\mathcal{S}$  is non-zero because we have limited ourselves to the situation where  $\mathbb{P}(\mathbf{H}^t) > 0$  and  $S^t = s$  with nonzero probability. We need only define  $F_+^t$  on the set of histories and injector locations that has some probability of occurring.

Next we consider  $\mathbb{P}(\mathcal{S}|\mathcal{C} \cap \mathbf{H}^{t-1})$ .

$$\mathbb{P}(\mathcal{S}|\mathcal{C} \cap \mathbf{H}^{t-1}) = \frac{\mathbb{P}(\mathcal{S} \cap \mathcal{C} \cap \mathcal{H}^{t-1})}{\mathbb{P}(\mathcal{C} \cap \mathbf{H}^{t-1})} \quad (4.3.5)$$

$$= \frac{\mathbb{P}(\mathcal{C}|\mathcal{S} \cap \mathbf{H}^{t-1})\mathbb{P}(\mathcal{S}|\mathbf{H}^{t-1})\mathbb{P}(\mathbf{H}^{t-1})}{\mathbb{P}(\mathcal{C}|\mathbf{H}^{t-1})\mathbb{P}(\mathbf{H}^{t-1})} \quad (4.3.6)$$

$$= \frac{\mathbb{P}(\mathcal{C}|\mathbf{H}^{t-1})f_-^t(s)}{\mathbb{P}(\mathcal{C}|\mathbf{H}^{t-1})} \quad (4.3.7)$$

$$= f_-^t(s) \quad (4.3.8)$$

Looking at Equation (4.3.6),  $\mathbb{P}(\mathcal{C}|\mathcal{S} \cap \mathbf{H}^{t-1}) = \mathbb{P}(\mathcal{C}|\mathbf{H}^{t-1})$  because the creature's location at time  $t$  is determined by  $\mathcal{H}^{t-1}$  and the creature's strategy. Furthermore,  $\mathbb{P}(\mathcal{S}|\mathbf{H}^{t-1}) = f_-^t(s)$  by definition of  $f_-^t(s)$ . Therefore

$$f_+^t(s) = \frac{\mathbb{P}(K_{c-s} = m_t)f_-^t(s)}{\mathbb{P}(\mathcal{M}|\mathcal{C} \cap \mathbf{H}^{t-1})}. \quad (4.3.9)$$

Turning our attention to the denominator of Equation: (4.3.9),

$$\mathbb{P}(\mathcal{M}|\mathcal{C} \cap \mathbf{H}^{t-1}) \quad (4.3.10)$$

$$= \sum_l \mathbb{P}(\mathcal{M} \cap S^t = l | \mathcal{C} \cap \mathbf{H}^{t-1}) \quad (4.3.11)$$

$$= \sum_l \frac{\mathbb{P}(\mathcal{M} \cap S^t = l \cap \mathcal{C} \cap \mathbf{H}^{t-1})}{\mathbb{P}(\mathcal{C} \cap \mathbf{H}^{t-1})} \quad (4.3.12)$$

$$= \sum_l \frac{\mathbb{P}(\mathcal{M}|S^t = l \cap \mathcal{C} \cap \mathbf{H}^{t-1}) \cdot \mathbb{P}(S^t = l | \mathcal{C} \cap \mathbf{H}^{t-1}) \cdot \mathbb{P}(\mathcal{C} \cap \mathbf{H}^{t-1})}{\mathbb{P}(\mathcal{C} \cap \mathbf{H}^{t-1})} \quad (4.3.13)$$

$$= \sum_l \mathbb{P}(K_{c-l} = m_t) f_-^t(l) \quad (4.3.14)$$

Equation (4.3.11) holds because the events  $S^t = l$  for all  $l$  partition the event space. Equation (4.3.14) follows for the same reasons that Equation (4.3.2) implies Equation (4.3.9). With these equations the creature can now calculate  $f_+^t$  using only  $m_t$ ,  $c_t$ , all  $\lambda_i$ 's, and  $f_-^t$ .

$$f_+^t(s) = \frac{\mathbb{P}(K_{c-s} = m_t) f_-^t(s)}{\sum_l \mathbb{P}(K_{c-l} = m_t) f_-^t(l)} \quad (4.3.15)$$

#### 4.4 The Injector's Next Location, Given a Full History

For this investigation, all movement algorithms for the injector had a set of probabilities,  $P = \{p_m : m \in \{0, \dots, n-1\}\}$ , where  $\sum_m p_m = 1$  and  $p_m$  is the probability of moving  $m$  steps to the right of its current location. Also, the injector would enact this algorithm at most once between successive observations.

To determine  $f_-^{t+1}$  the creature needs to calculate the following for each  $s \in \{0, \dots, N-1\}$ .

**Proposition 4.1.**  $F_-^{t+1}$  is the convolution of  $F_+^t$  with  $P$ .

*Proof.* The probability that the injector is at location  $l$  after moving is the probability the injector was  $m$  steps to the left of  $l$  and then moved  $m$  steps to the right, summed over each  $m \in \{0, \dots, N-1\}$ .

$$F_-^{t+1}(l) = \sum_m F_+^t(l-m) \cdot p_m, \quad (4.4.1)$$

where  $F_+^t(l-m)$  is the probability that the injector was  $m$  steps to left of  $l$ . Equivalently,  $F_-^{t+1}$  is a convolution of the probability mass functions  $F_+^t$  and  $P$ . It is important to note that this equation is only valid if  $p_m$  is independent of the injector's location. □

With Equation: (4.4.1), the creature can calculate  $f_-^{t+1}$  using only  $f_+^t$  and  $P$ . Let us consider the algorithms the different injector movement algorithms (See Section: 2.2.3):



1.  $P_0$

In this case the injector is stationary.

$$\begin{aligned} F_-^{t+1}(l) &= \sum_m F_+^t(l-m) \cdot p_m \\ &= F_+^t(l) \end{aligned}$$

$P_0$  is a delta function, which is the identity element for convolutions. Therefore

$$F_-^{t+1} = F_+^t * P_0 = f_+^t.$$

2.  $P_{uniform}$

In this case the injector is equally likely to move anywhere and

$$\begin{aligned} F_-^{t+1}(l) &= \sum_m F_+^t(l-m) \cdot p_m \\ &= \frac{1}{N} \sum_m F_+^t(l-m) \\ &= \frac{1}{N} \end{aligned}$$

This means that any prior knowledge of the injector's location is erased each time the injector moves.

$P_{uniform}$  is the constant function for convolutions and therefore  $F_-^{t+1} = F_+^t * P_{uniform} = P_{uniform}$ .

3.  $P_x$

This is the standard case.

$$\begin{aligned} F_-^{t+1}(l) &= \sum_m F_+^t(l-m) \cdot p_m \\ &= x \cdot F_+^t(l-1) + (1-2x) \cdot F_+^t(l) + x \cdot F_+^t(l+1) \end{aligned}$$

## 5 Injection Rates in the Limit

This section will analyze some of the properties of  $F_+^t$  as the injection rate becomes large. The function  $F_+^t(l)$  is the probability that the injector is at location  $l$  given the creature's full history up to and including time  $t$ . Its form is given by Equation (4.3.15). Consider the initial observation made by the creature. For this thesis, the creature has no information about the injector's location before the first observation. This means that  $f_-^0(l) = 1/N$ .

For this section we fix the creature's location to be  $c$  and the source's location to be  $s$ . The distance from the creature to the injector,  $d$ , is the minimum of  $c - s$  and  $s - c$ .<sup>4</sup>

**Theorem 5.1.** *As  $\gamma \rightarrow \infty$ , the expected value of  $F_+^0(l)$  conditioned on  $M^0$  goes to 0 for all  $l \neq c \pm d$ .*

*Proof.* Let  $d'$  be distance between the creature and a location  $l \neq c \pm d$ . (For a diagram of the locations and distances referred to in this proof consult Figure 5.1.) The number of molecules observed at time  $t$ , namely  $M^0$ , is just  $K_d$  (a Poisson random variable with mean  $\lambda_d$ ). We will write  $\mathbb{E}_M$  for expectation with respect to the distribution of  $M^0$  (equivalently,  $K_d$ ). Using Equation (4.3.15)

$$\mathbb{E}_M(F_+^0(l)) = \sum_{k=0}^{\infty} \mathbb{P}(K_d = k) \cdot \frac{\mathbb{P}(K_{d'} = k) f_-^0(l)}{\sum_n \mathbb{P}(K_{c-n} = k) f_-^0(n)}. \quad (5.0.2)$$

We will partition this sum into two different sums and then find a bound for each.

Let  $\epsilon > 0$  and consider the set  $A$  of possible molecule counts that are close to  $\lambda_d$ ,

$$A := \left\{ k \in \mathbb{N} : \lambda_d - \frac{|\lambda_d - \lambda_{d'}|}{2} \leq k \leq \lambda_d + \frac{|\lambda_d - \lambda_{d'}|}{2} \right\}.$$

Because the difference  $\lambda_d - \lambda_{d'}$  grows linearly with  $\gamma$  (See Lemma 3.11), for large  $\gamma$  we will obviously have  $A$  nonempty as  $\gamma \rightarrow \infty$ .

We will find a  $\gamma'$  such that for all  $\gamma > \gamma'$

---

<sup>4</sup>The addition and subtraction of locations occurs in  $\mathbb{Z}_N$ .

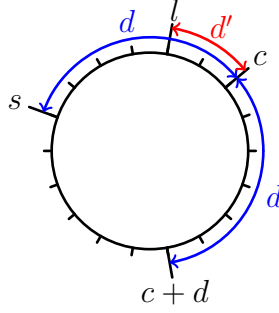


Figure 5.1: A diagram of locations and distances for Theorem 5.1. The creature is at location  $c$  and the injector is at location  $s$ , which is a distance  $d = |c - s|$  from the creature. Location  $l$ , which is distance  $d'$  from the creature, is a generic location on the ring where  $d \neq d'$ . It is possible that the true source location  $s$  could be  $c + d$  instead of  $c - d$ . The proof of Theorem 5.1 is equally valid in either case.

$$\sum_{k \in A} \mathbb{P}(K_d = k) \cdot \frac{\mathbb{P}(K_{d'} = k) f_-^0(l)}{\sum_n \mathbb{P}(K_{c-n} = k) f_-^0(n)} < \epsilon/2 \quad (5.0.3)$$

$$\sum_{k \notin A} \mathbb{P}(K_d = k) \cdot \frac{\mathbb{P}(K_{d'} = k) f_-^0(l)}{\sum_n \mathbb{P}(K_{c-n} = k) f_-^0(n)} < \epsilon/2 \quad (5.0.4)$$

Looking at Equation 5.0.3

$$\frac{1}{f_-^0(s)} \sum_{k \in A} \mathbb{P}(K_{d'} = k) > \sum_{k \in A} \mathbb{P}(K_d = k) \cdot \frac{\mathbb{P}(K_{d'} = k) f_-^0(l)}{\sum_n \mathbb{P}(K_{c-n} = k) f_-^0(n)} \quad (5.0.5)$$

because

$$\frac{1}{f_-^0(s)} \geq \mathbb{P}(K_d = k) \cdot \frac{f_-^0(l)}{\sum_n \mathbb{P}(K_{c-n} = k) f_-^0(n)}.$$

The term  $f_-^0(s)$  is equal to  $1/N$ , therefore  $\frac{1}{f_-^0(s)} = N$ . For Equation 5.0.4

$$\sum_{k \notin A} \mathbb{P}(K_d = k) > \sum_{k \notin A} \mathbb{P}(K_d = k) \cdot \frac{\mathbb{P}(K_{d'} = k) f_-^0(l)}{\sum_n \mathbb{P}(K_{c-n} = k) f_-^0(n)} \quad (5.0.6)$$

because

$$1 \geq \frac{\mathbb{P}(K_{d'} = k) f_-^0(l)}{\sum_n \mathbb{P}(K_{c-n} = k) f_-^0(n)}.$$

In order to find a  $\gamma'$  that bounds Equations (5.0.5) and (5.0.6), we will make use of the following facts.  $K_d$  has mean  $\lambda_d$  and standard deviation  $\sigma_d = \sqrt{\lambda_d}$ . From Corollary 3.13 and Lemma 3.12

$$\begin{aligned}\lambda_d &= \gamma \lambda_d(1) \\ \sigma_d &= \sqrt{\gamma \lambda_d(1)}\end{aligned}$$

Chebysheff's inequality states

$$\mathbb{P}(|K_d - \lambda_d| \geq s\sigma_d) < 1/s^2. \quad (5.0.7)$$

Equivalent statements can be made for  $K_{d'}$ . Furthermore, from Lemma 3.10 it is true that  $\lambda_d \neq \lambda_{d'}$  because  $d \neq \pm d'$  and from Lemma 3.9  $\lambda_d, \lambda_{d'} > 0$ .

The set  $B := \{k \in \mathbb{N} : \lambda_{d'} - \frac{|\lambda_d - \lambda_{d'}|}{2} \leq k \leq \lambda_{d'} + \frac{|\lambda_d - \lambda_{d'}|}{2}\}$ , is the set of possible molecule counts close to  $\lambda_{d'}$ . Note that  $A$  is a subset of  $B^c$  so  $\sum_{k \notin B} \mathbb{P}(K_{d'} = k) \geq \sum_{k \in A} \mathbb{P}(K_{d'} = k)$ . Therefore,

$$\sum_{k \notin B} \mathbb{P}(K_{d'} = k) = \mathbb{P}\left(|K_{d'} - \lambda_{d'}| \geq \frac{|\lambda_d - \lambda_{d'}|}{2}\right).$$

Let  $\theta = \frac{\lambda_d}{\lambda_{d'}} = \frac{\lambda_d(1)}{\lambda_{d'}(1)}$ . We can see that  $\theta \neq 1$  because  $\lambda_d \neq \lambda_{d'}$  and  $\theta \notin \{0, \infty\}$  because  $\lambda_d, \lambda_{d'} > 0$  (See Lemma 3.9). Consequently,

$$\begin{aligned}\mathbb{P}\left(|K_{d'} - \lambda_{d'}| \geq \frac{|\lambda_d - \lambda_{d'}|}{2}\right) &= \mathbb{P}\left(|K_{d'} - \lambda_{d'}| \geq \frac{|\theta - 1|}{2} \lambda_{d'}\right) \\ &= \mathbb{P}\left(|K_{d'} - \lambda_{d'}| \geq \frac{|\theta - 1|}{2} \sigma_{d'}^2\right) \\ &\leq \left(\frac{2}{|\theta - 1| \sigma_{d'}}\right)^2 \\ &= \left(\frac{2}{|\theta - 1|}\right)^2 \frac{1}{\gamma \lambda_{d'}(1)}.\end{aligned} \quad (5.0.8)$$

The inequality holds due to Equation (5.0.7). Let

$$\gamma'_1 > N \left(\frac{2}{|\theta - 1|}\right)^2 \frac{2}{\epsilon \lambda_{d'}(1)}. \quad (5.0.9)$$

From Equation (5.0.8), we can see that whenever  $\gamma > \gamma'_1$ ,  $\frac{1}{f^0(s)} \sum_{k \in A} \mathbb{P}(K_{d'} = k) < \epsilon/2$ . We can let  $\gamma$  tend to infinity because  $\frac{1}{f^0(s)} = 1/N$  is independent of  $\gamma$  as are  $\lambda_{d'}(1)$  and  $\theta = \frac{\lambda_d(1)}{\lambda_{d'}(1)}$ .

A similar line of logic applied to Equation (5.0.3) discovers there is a

$$\gamma'_2 > \left( \frac{2}{|\theta^{-1} - 1|} \right)^2 \frac{2}{\epsilon \lambda_d(1)},$$

such that for all  $\gamma > \gamma'_2$ ,  $\sum_{k \notin A} \mathbb{P}(K_d = k) < \epsilon/2$ . Let  $\gamma' = \max\{\gamma'_1, \gamma'_2\}$ . Now that we have established Equations (5.0.3) and (5.0.4), we can conclude that when  $l \neq \pm d$  we have

$$\mathbb{E}_M[F_+^0(l)] \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

□

**Remark 5.2.** *Theorem 5.1 implies that as  $\gamma \rightarrow \infty$ , the expected probability that the injector is at any location  $l$  that is closer to or further from the creature than the injector goes to 0.*

Having described the expected value of  $F_+^0(l)$  when  $l \neq c \pm d$ , it remains to describe  $\mathbb{E}_M(F_+^0(c+d))$  and  $\mathbb{E}_M(F_+^0(c-d))$ .

**Corollary 5.3.** *As  $\gamma \rightarrow \infty$ , when  $d \in \{0, N/2\}$ ,*

$$\mathbb{E}_M(F_+^0(c+d)) \rightarrow 1.$$

*For all other  $d$ ,*

$$\mathbb{E}_M(F_+^0(c+d)) \rightarrow \frac{F_-^0(c+d)}{F_-^0(c+d) + F_-^0(c-d)}$$

*and*

$$\mathbb{E}_M(F_+^0(c-d)) \rightarrow \frac{F_-^0(c-d)}{F_-^0(c+d) + F_-^0(c-d)}.$$

*Proof.* From Theorem 5.1,  $\lim_{\gamma \rightarrow \infty} \mathbb{E}_M F_+^0(l) \rightarrow 0$  for all  $l \neq c \pm d$ . Because  $F_+^0$  is a random variable describing a probability mass function,  $\lim_{\gamma \rightarrow \infty} \mathbb{E}_M \sum_l F_+^0(l) = 1$ .

Consider the case when  $d \in \{0, N/2\}$ . This means  $c + d = c - d$ . So

$$\begin{aligned}
1 &= \lim_{\gamma \rightarrow \infty} \mathbb{E}_M \sum_l (F_+^0(l)) \\
&= \lim_{\gamma \rightarrow \infty} (\mathbb{E}_M F_+^0(c + d)) + \sum_{l \neq c+d} \lim_{\gamma \rightarrow \infty} (\mathbb{E}_M F_+^0(l)) \\
&= \lim_{\gamma \rightarrow \infty} (\mathbb{E}_M F_+^0(c + d)).
\end{aligned}$$

For the case where  $d \notin \{0, N/2\}$

$$1 = \lim_{\gamma \rightarrow \infty} \mathbb{E}_M(F_+^0(c + d)) + \lim_{\gamma \rightarrow \infty} \mathbb{E}_M(F_+^0(c - d)).$$

Turning our attention to the expected value of  $F_+^0(c + d)$  conditioned on  $M^0$ ,

$$\begin{aligned}
\mathbb{E}_M(F_+^0(c + d)) &= \sum_{k=0}^{\infty} \mathbb{P}(K_d = k) \cdot \frac{\mathbb{P}(K_d = k) F_-^0(c + d)}{\sum_n \mathbb{P}(K_{c-n} = k) F_-^0(n)} \\
&\leq \sum_{k=0}^{\infty} \mathbb{P}(K_d = k) \cdot \frac{\mathbb{P}(K_d = k) F_-^0(c + d)}{\mathbb{P}(K_d = k) (F_-^0(c + d) + F_-^0(c - d))} \\
&= \frac{F_-^0(c + d)}{F_-^0(c + d) + F_-^0(c - d)} \sum_{k=0}^{\infty} \mathbb{P}(K_d = k) \\
&= \frac{F_-^0(c + d)}{F_-^0(c + d) + F_-^0(c - d)}. \tag{5.0.10}
\end{aligned}$$

The inequality holds because

$$\sum_n \mathbb{P}(K_{c-n} = k) F_-^0(n) \geq \mathbb{P}(K_d = k) F_-^0(c + d) + \mathbb{P}(K_{-d} = k) F_-^0(c - d)$$

and  $K_d = K_{-d}$ . Similarly,

$$\mathbb{E}_M(F_+^0(c - d)) \leq \frac{F_-^0(c - d)}{F_-^0(c + d) + F_-^0(c - d)} \tag{5.0.11}$$

Therefore,

$$\begin{aligned}
1 &= \lim_{\gamma \rightarrow \infty} \mathbb{E}_M(F_+^0(c + d)) + \lim_{\gamma \rightarrow \infty} \mathbb{E}_M(F_+^0(c - d)) \\
&\leq \frac{F_-^0(c + d)}{F_-^0(c + d) + F_-^0(c - d)} + \frac{F_-^0(c - d)}{F_-^0(c + d) + F_-^0(c - d)} \\
&= 1.
\end{aligned}$$

This means the  $\leq$  is an equality. Therefore

$$\lim_{\gamma \rightarrow \infty} \mathbb{E}_M(F_+(c+d)) = \frac{F_-^0(c+d)}{F_-^0(c+d) + F_-^0(c-d)}$$

and

$$\lim_{\gamma \rightarrow \infty} \mathbb{E}_M(F_+(c-d)) = \frac{F_-^0(c-d)}{F_-^0(c+d) + F_-^0(c-d)}.$$

□

These results suggest that for sufficiently large  $\gamma$ , the behavior of the model system becomes similar to that of a system in which the creature can determine its distance to the injector on each time step *with certainty*. We will consider such a “distance certain” version of our model system in Section 6. Because the movement of the injector under the standard movement algorithm is stochastic, uncertainty about its future location persists even if the distances to its past locations could be known with certainty.

The sense in which the system described thus far “becomes similar to” the simpler distance-certain system for large  $\gamma$  is technically subtle. Developing a precise formulation that captures this heuristic observation, and proofs of additional results in this direction, will await future work. In the next Section we will proceed by making the strong assumption that for sufficiently large  $\gamma$  the system behaves as if all distances could be determined with certainty.

## 6 Coarse Graining for a Simplified Model

### 6.1 Introduction

In Section 5 we explored the probabilistic properties of the injector's location given a full history as the injection rate became large. For the rest of this thesis, we will continue to assume that the molecule movement algorithm is described by  $Q_y$  and the injector movement algorithm is described by  $P_x$ . We also simplify the model (in a manner suggested by Section 5) and consider a system where an observation of molecules almost surely identifies the distance from the injector to the creature.

A creature strategy is a set of rules describing how the creature moves when it is in any given situation. A creature's "situation" is its entire history. Because our creatures are only concerned with the amount of food they get or how much information they can gain about the injector's location, we developed a method of turning a history into a p.m.f. of the injector's next location.

Let  $\mathcal{J}_+$  be the set of  $f_+ \in \mathcal{P}(\mathbb{Z}_N)$  such that either there is a location  $e$  for which  $f_+(e) = 1$ , or else there is a location  $e$  and a distance  $d > 0$  such that  $f_+(e+d) + f_+(e-d) = 1$ . In our simplified model the support of  $F_+^t$  is a subset of  $\mathcal{J}_+$ . In this section we will classify the support of the  $F_+^t$ 's into "information states" and then do the same for the support of the  $F_-^t$ 's. We define an information state to be a partition of the support of  $F_+^t$  or the support of  $F_-^t$ . The rest of this section creates a partitioning of these supports.

### 6.2 Information States for $F_+^t$

The set  $\mathcal{J}_+$  is the set of all  $f_+ \in \mathcal{P}(\mathbb{Z}_N)$  where the creature almost surely knows the injector is at some location  $e$ , or where the creature knows the injector is almost surely some distance  $d$  away from some location  $e$ . As suggested by the discussion in the previous Section, we will assume that the support of  $F_+^t$  is a subset of  $\mathcal{J}_+$ . An information state is a set of  $f_+ \in \mathcal{J}_+$  with specific properties. The sets are mutually



disjoint and their union is  $\mathcal{J}_+$ .

Let the information state  $\mathcal{J}_+^0$  be the set of  $f_+ \in \mathcal{P}(\mathbb{Z}_N)$  such that there exists an  $l$  where  $f_+(l) = 1$ . This state describes the situation where the creature knows exactly where the injector is located at time  $t$ . This can occur in two ways.

First, as shown in Theorem 5.1 the creature could be 0 or  $N/2$  steps away from the injector. Although it is reasonable that a creature that lands on the injector would be able to deduce the injector's location, it is disconcerting to see that landing on the opposite side of the ring also allows the creature deduce the injector's location. This irregularity will be seen in some of the other information states; however, it is not a significant blow against the model when the creature stays away from the far side of the ring. As seen in Section 7 the creature strategies studied in this thesis obey that rule.

Second, from Corollary 5.3 if  $f_-^t(c+d) = 0$  then  $f_-^t(c-d) = 1$  and vice versa.<sup>5</sup> Essentially this describes the situation where the creature knows it is  $d$  steps to the left or right based on the observation  $M^t$  and using its knowledge of the history, eliminates one of those possibilities.

The information state  $\mathcal{J}_+^0$  covers the situation where the creature knows the injector's exact location. Assuming the creature always know its distance from the injector, the only  $f_+^t$  left to describe are those created by the situation where  $S^t = c \pm d$ ,  $0 < d < N/2$ , and both  $f_+^t(c+d) > 0$  and  $f_+^t(c-d) > 0$ .

The information state  $\mathcal{J}_+^1 = \{f_+ \in \mathcal{P}(\mathbb{Z}_N) : f_+(c+1) = f_+(c-1) = 1/2 \text{ or } f_+(c'+1) = f_+(c'-1) = 1/2\}$ . This state describes the situation where the creature knows the injector is equally likely to be one step to the right or left of the creature. It also includes the situation where the creature learns the injector is  $N/2 - 1$  steps away.

The information state  $\mathcal{J}_+^2 := \{f_+ \in \mathcal{P}(\mathbb{Z}_N) : f_+(c+2) = f_+(c-2) = 1/2 \text{ or } f_+(c'+2) = f_+(c'-2) = 1/2\}$ . This state describes the situation where the creature knows

---

<sup>5</sup>For brevity we will refer to the creature's position  $C^t$  as " $c$ ", and the opposite position on the ring as  $c' = c + N/2$ .

the injector is equally likely to be two steps to the right or left of the creature. It also includes the situation where the creature learns the injector is  $N/2 - 2$  steps away.

The information state  $\mathcal{J}_+^{3+} := \{f_+ \in \mathcal{P}(\mathbb{Z}_N) : \text{there exists a } d \text{ with } 2 < d < N/2 - 2 \text{ such that } f_+(c+d) = f_+(c-d) = 1/2\}$ . This state describes the situation where the creature knows the injector is equally likely to be three or more steps to the right or left of the creature but not covered by  $\mathcal{J}_+^1$  or  $\mathcal{J}_+^2$ .

Having covered all situations where the injector is equally likely to be  $l$  steps to the right or the left of the injector, there only remain the cases where  $S^t = c \pm d$ ,  $0 < d < N/2$ ,  $f_+^t(c+d) > 0$ ,  $f_+^t(c-d) > 0$  and  $f_+^t(c+d) \neq f_+^t(c-d)$ . Thus we introduce the information state  $\mathcal{J}_+^I := \{f_+ \in \mathcal{P}(\mathbb{Z}_N) : f_+(c+d) \neq f_+(c-d) \text{ and } d > 0, f_+(c+d) > 0, f_+(c-d) > 0\}$ . We refer to such an information state as *imbalanced*. Figure 6.2 illustrates several information states.

### 6.3 Information States for $F_-^t$

The set  $\mathcal{J}_- := \{f_- \in \mathcal{P}(\mathbb{Z}_N) : f_- = f_+ * P_x \text{ with } f_+ \in \mathcal{J}_+\}$  contains the support of  $F_-^t$  in the simplified model. Having partitioned  $\mathcal{J}_+$  it is straight forward to classify  $\mathcal{J}_-$ . From Proposition 4.1  $F_-^{t+1}$  is the convolution of  $F_+^t$  with  $P_x$ . The only  $F_-^t$  for which this does not apply is  $F_-^0$ . Let  $\mathcal{J}_-^* = \{f_-^0\}$ . As discussed in Section 2.2.5,  $f_-^0(l) = 1/N$  for all  $l$ . For each information state in  $\mathcal{J}_+$  there will be a corresponding information state for  $\mathcal{J}_-$ . In this section we will explore the properties of the  $\mathcal{J}_-$  information states. As we shall show, the properties of these new states are mutually exclusive. Furthermore, as  $f_-^t$  is derived from  $f_+^t$ , all  $f_-^t$ 's that occur will exist in one of these new information states.

For each  $f_+$  in  $\mathcal{J}_+$  there exists some location  $e$  such that  $f_+(e) = 1$  or  $f_+(e+l) + f_+(e-l) = 1$  with  $0 < l \leq N/4$ . Specifically,  $e \in \{c, c'\}$ . Consequently, if  $f_+ * P_x$  is invertible,<sup>6</sup> each  $f_- \in \mathcal{J}_-$  can also be associated with  $e$ . For the rest of this thesis let  $e$  be that location for the  $f_+$  and/or  $f_-$  under discussion.

---

<sup>6</sup>Convolutions are not generally invertible, but specific convolutions can be and in this case are.

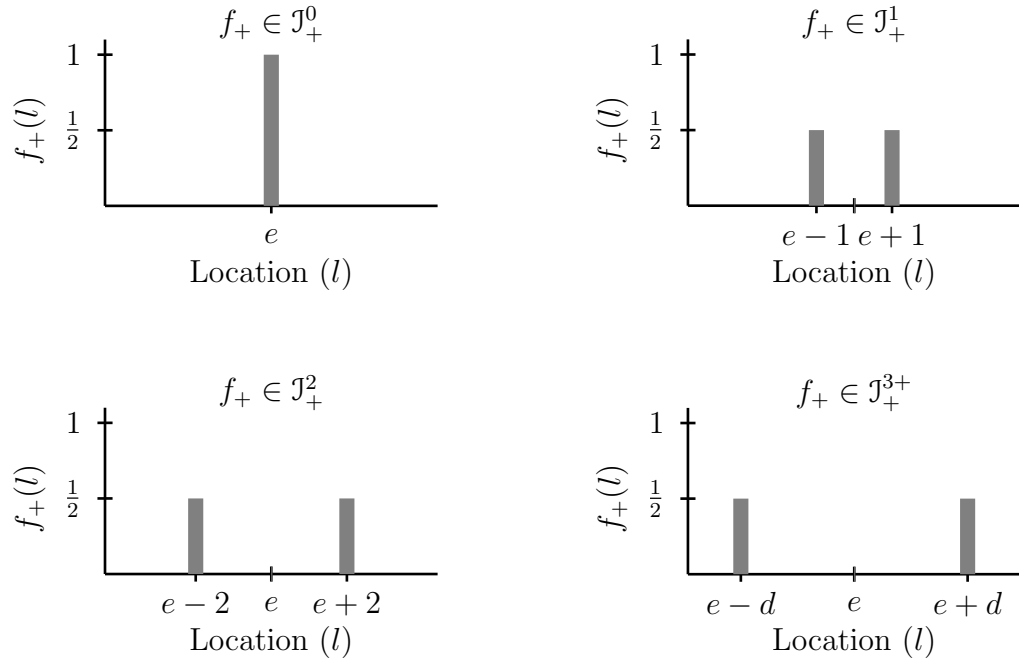


Figure 6.1: The primary information states for the support of  $F_+^t$  (in the distance-certain model) namely  $\mathcal{J}_+$ . The function  $f_+$  is a probability mass function on  $\mathbb{Z}_N$ . For every  $f_+$  in  $\mathcal{J}_+^0$  there is a location  $e$  such that  $f_+(e) = 1$ . For the other information states there are two locations equidistant from a location  $e$  where the injector is equally likely to be; furthermore the injector is at any other location with probability 0. For  $f_+$  to be in  $\mathcal{J}_+^{3+}$ ,  $d$  must be greater than 2 but less than or equal to  $N/4$ .

### 6.3.1 The Information State $\mathcal{J}_-^0$

The information state  $\mathcal{J}_-^0 := \{f_- \in \mathcal{P}(\mathbb{Z}_N) : f_- = f_+ * P_x \text{ with } f_+ \in \mathcal{J}_+^0\}$ . Let us look at some of the properties of  $f_- \in \mathcal{J}_-^0$ . There exists  $f_+ \in \mathcal{J}_+^0$  such that  $f_- = f_+ * P_x$ . Furthermore,  $f_+(e) = 1$ . Therefore,

$$f_-(l) = \begin{cases} x & l = e \pm 1 \\ 1 - 2x & l = e \\ 0 & \text{otherwise.} \end{cases} \quad (6.3.1)$$

Having characterized  $f_-$  we now calculate some of its properties. From Equations (6.3.1) and (1.1.1), when  $f_- \in \mathcal{J}_-^0$ , the entropy of  $f_-$  is

$$H(f_-) = -2x \log_2(x) - (1 - 2x) \log_2(1 - 2x) \quad (6.3.2)$$

Let us turn our attention to what happens when the creature chooses to move to certain locations. It is important to remember that when the creature decides where to move to it has access to the probability distribution of the injector's next location, namely  $f_-^{t+1}$ . When  $f_-^{t+1} \in \mathcal{J}_-^0$ , the creature has a set of expected results for moving to different locations on the ring. See Figures 6.2 and 6.3 for a visual representation.

1. Location  $C^{t+1} = c_+ \in \{e \pm 1\}$ .

Due to the symmetry of  $\hat{\lambda}$  and  $f_- \in \mathcal{J}_-^0$ , the expected amount of food and expected distance from the injector during the next time step for  $e + 1$  is the same as  $e - 1$ . Therefore we discuss them together. Without loss of generality assume  $c_+ = e + 1$ .

The expected food for a creature that moves to  $c_+$  given  $f_-^t \in \mathcal{J}_-^0$  is the sum over  $l$  of the probability the injector is at  $l$  times the expected amount of food the creature will get when it is at  $c_+$  and the injector is at  $l$ .

$$\begin{aligned} \mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+-l}) \\ &= x\lambda_0 + (1 - 2x)\lambda_1 + x\lambda_2 \end{aligned} \quad (6.3.3)$$

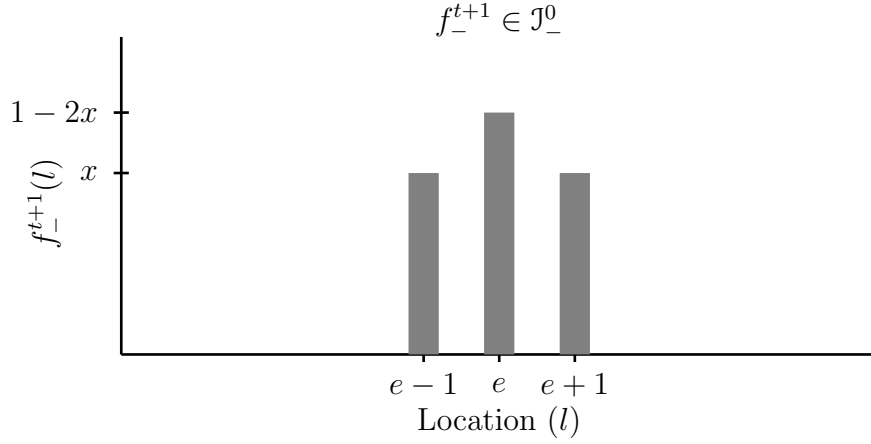


Figure 6.2: A typical probability mass function  $f_-$  in the information state  $\mathcal{J}_-^0$ . The creature is in the information state  $\mathcal{J}_-^0$  at time  $t$  if and only if  $F_+^t \in \mathcal{J}_-^0$ .

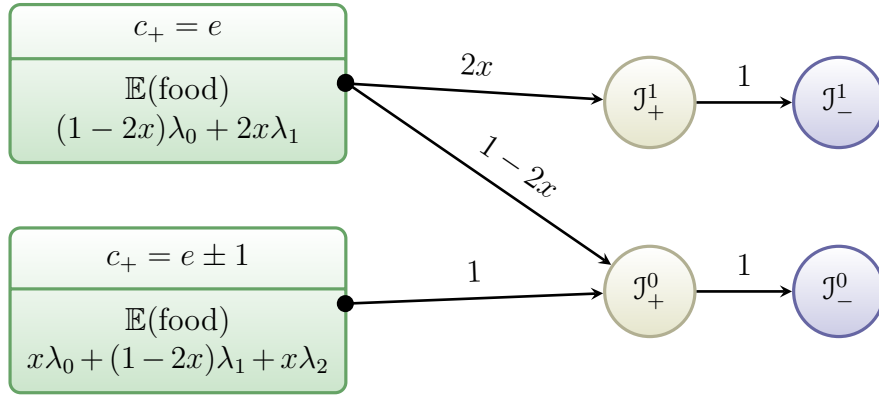


Figure 6.3: Transitions from  $\mathcal{J}_-^0$ . This figure states the expected food at each choice for the creature's next location ( $c_+$ ) when in the information state  $\mathcal{J}_-^0$ . It also depicts the transition probabilities to other information states based on  $c_+$ . By definition, if the creature is in  $\mathcal{J}_+^0$  or  $\mathcal{J}_+^1$  at time  $t$  the creature will be in  $\mathcal{J}_-^0$  or  $\mathcal{J}_-^1$  at time  $t+1$  (respectively). Future graphs will not depict the intermediate information states.

It is also possible to predict what  $f_+^{t+1}$  and hence  $f_-^{t+2}$  are from the creature's choice of location. If the creature goes to  $c_+$  there is a probability  $x$  that the injector will also go to  $c_+$ . If that happens the creature will almost surely know the injector is 0 steps away from location  $c_+$  (Corollary 5.3). This means that  $f_+^{t+1} \in \mathcal{J}_+^0$  and  $f_-^{t+2} \in \mathcal{J}_-^0$ .

There is a probability  $1 - 2x$  that the injector will go to  $c_+ - 1 = e$ . If that happens the creature will almost surely know the injector is 1 step away from location  $c_+$  (Corollary 5.3) but the probability that the injector is at  $c_+ + 1 = e + 2$  is 0. Therefore,  $f_+^{t+1} \in \mathcal{J}_+^0$  and  $f_-^{t+2} \in \mathcal{J}_-^0$ . Similarly, if the injector goes to  $c_+ - 2 = e - 1$ , then  $f_+^{t+1} \in \mathcal{J}_+^0$  and  $f_-^{t+2} \in \mathcal{J}_-^0$ .

Therefore we have shown if  $f_-^{t+1} \in \mathcal{J}_-^0$  and the creature goes to  $e \pm 1$ , then  $f_-^{t+2}$  is almost surely in  $\mathcal{J}_-^0$ .

2. Location  $C^{t+1} = c_+ = e$ .

If the creature goes to location  $e$  the expected food is

$$\begin{aligned} \mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+ - l}) \\ &= (1 - 2x)\lambda_0 + 2x\lambda_1 \end{aligned} \tag{6.3.4}$$

The probability that a creature's  $f_+^{t+1}$  is in each partition of  $\mathcal{J}_+$  and hence that  $f_-^{t+2}$  in  $\mathcal{J}_+$  is listed in Figure 6.2. When the creature goes to location  $e$  there is a probability  $1 - 2x$  that the injector also moves to location  $e$ . In that case the next observation of molecules tells the creature that the injector is almost surely 0 steps away. This means the injector almost surely is at location  $e$  and  $f_+^{t+1} \in \mathcal{J}_+^0$ . If  $f_+^{t+1} \in \mathcal{J}_+^0$ , then  $f_-^{t+2}$  in  $\mathcal{J}_+^1$ .

There is a probability  $x$  that the injector will be at  $e + 1$  and the same for  $e - 1$ . If the creature is at location  $e$ , the next observation will indicate the injector is almost surely 1 step away. As it is equally likely for for the injector to be to right or left,  $f_+^{t+1} \in \mathcal{J}_+^1$ . If  $f_+^{t+1} \in \mathcal{J}_+^1$ , then  $f_-^{t+2}$  in  $\mathcal{J}_+^1$ .

### 6.3.2 The Information State $\mathcal{J}_-^1$

The information state  $\mathcal{J}_-^1 := \{f_- \in \mathcal{P}(\mathbb{Z}_N) : f_- = f_+ * P_x \text{ with } f_+ \in \mathcal{J}_+^1\}$ . Let us look at some of the properties of  $f_- \in \mathcal{J}_-^1$ . There exists  $f_+ \in \mathcal{J}_+^1$  such that  $f_- = f_+ * P_x$ . Furthermore,  $f_+(e-1) = f_+(e+1) = 1/2$ . Therefore,

$$f_-(l) = \begin{cases} \frac{x}{2} & l = e \pm 2 \\ \frac{1-2x}{2} & l = e \pm 1 \\ x & l = e \\ 0 & \text{otherwise.} \end{cases} \quad (6.3.5)$$

From Equations (6.3.5) and (1.1.1), when  $f'_- \in \mathcal{J}_-^1$ , the entropy of  $f'_-$  is

$$H(f'_-) = -x \log_2(x/2) - (1-2x) \log_2((1-2x)/2) - x \log_2(x) \quad (6.3.6)$$

**Lemma 6.1.** *The entropy of any  $f_- \in \mathcal{J}_-^1$  is greater than the entropy of any  $f'_- \in \mathcal{J}_-^0$ .*

*Proof.* From Equations (6.3.2) and (6.3.6), if  $f_- \in \mathcal{J}_-^1$  and  $f'_- \in \mathcal{J}_-^0$ ,

$$\begin{aligned} H(f_-) - H(f'_-) &= -x \log_2(x/2) - (1-2x) \log_2((1-2x)/2) \\ &\quad + x \log_2(x) + (1-2x) \log_2(1-2x) \\ &= x(\log_2(x) - \log_2(x/2)) \\ &\quad + (1-2x)(\log_2(1-2x) - \log_2((1-2x)/2)) \\ &= x \log_2(2) + (1-2x)(\log_2(2)) \\ &= (1-x) \log_2(2) \\ &> 0 \end{aligned}$$

The third equality holds because  $0 < x < 1/2$ . □

Let us turn our attention to what happens when the creature chooses to move to certain locations. It is important to remember that when the creature decides where to move to it has access to the probability distribution of the injector's next location, namely  $f_-^{t+1}$ . When  $f_-^{t+1} \in \mathcal{J}_-^1$ , the creature has a set of expected results for moving to different locations on the ring. See Figures 6.4 and 6.5 for a visual representation.

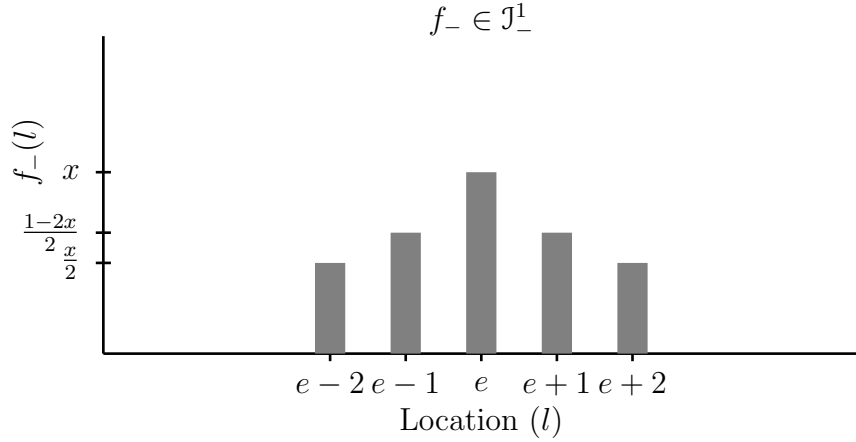


Figure 6.4: A typical probability mass function  $f_-$  in the information state  $\mathcal{J}_-^1$ . The creature is in the information state  $\mathcal{J}_-^1$  at time  $t$  if and only if  $F_+^t \in \mathcal{J}_-^1$ . The relative heights depicted hold for  $1/4 < x < 1/3$ . In this range we have  $x > (1 - 2x)/2 > x/2$ .

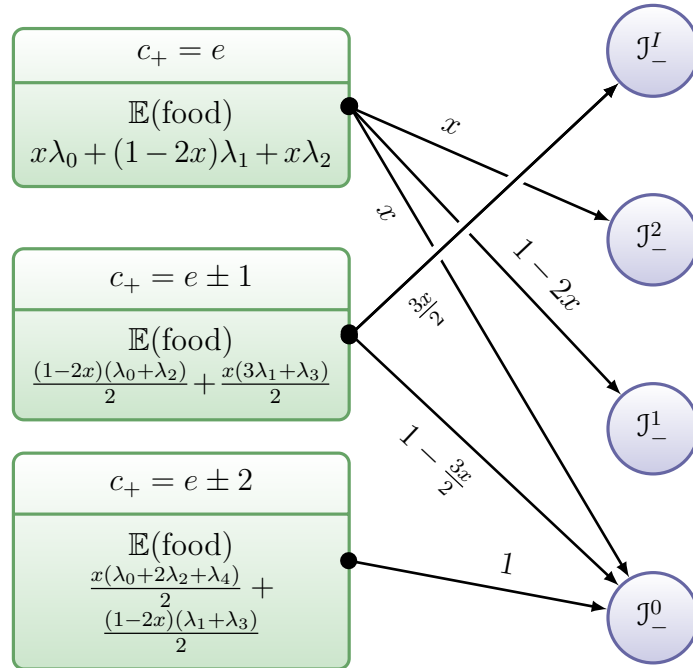


Figure 6.5: Transitions from  $\mathcal{J}_-^1$ . This figure states the expected food at each choice for the creature's next location ( $c_+$ ) when in the information state  $\mathcal{J}_-^1$ . It also depicts the transition probabilities to other information states based on  $c_+$ .



1. Location  $C^{t+1} = c_+ \in \{e \pm 2\}$ . Without loss of generality assume  $c_+ = e + 2$ .

The expected food for a creature that moves to  $c_+$  given  $f_-^t \in \mathcal{J}_-^0$  is the sum over  $l$  of the probability the injector is at  $l$  times the expected amount of food the creature will get when it is at  $c_+$  and the injector is at  $l$ .

$$\begin{aligned}\mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+-l}) \\ &= \frac{x(\lambda_0 + 2\lambda_2 + \lambda_4)}{2} + \frac{(1-2x)(\lambda_1 + \lambda_3)}{2}\end{aligned}\quad (6.3.7)$$

Using the same arguments present in Section 6.3.1, if the creature goes to  $e \pm 2$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1$$

2. Location  $C^{t+1} = c_+ \in \{e \pm 1\}$ .

Without loss of generality assume  $c_+ = e + 1$ .

$$\begin{aligned}\mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+-l}) \\ &= \frac{(1-2x)(\lambda_0 + \lambda_2)}{2} + \frac{x(3\lambda_1 + \lambda_3)}{2}\end{aligned}\quad (6.3.8)$$

Using the same arguments present in Section 6.3.1, if the creature goes to  $e \pm 1$

$$\begin{aligned}\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^I) &= \frac{3x}{2} \\ \mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) &= 1 - \frac{3x}{2}\end{aligned}$$

3. Location  $C^{t+1} = c_+ = e$ .

$$\begin{aligned}\mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+-l}) \\ &= \frac{(1-2x)\lambda_0}{2} + \frac{3x\lambda_1}{2} + \frac{(1-2x)\lambda_2}{2} + \frac{x\lambda_3}{2}\end{aligned}\quad (6.3.9)$$

Using the same arguments present in Section 6.3.1, if the creature goes to  $e$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^2) = x$$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = 1 - 2x$$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = x$$

### 6.3.3 The Information State $\mathcal{J}_-^2$

The information state  $\mathcal{J}_-^2 := \{f_- \in \mathcal{P}(\mathbb{Z}_N) : f_- = f_+ * P_x \text{ with } f_+ \in \mathcal{J}_+^2\}$ . Let us look at some of the properties of  $f_- \in \mathcal{J}_-^2$ . There exists  $f_+ \in \mathcal{J}_+^2$  such that  $f_- = f_+ * P_x$ . Furthermore,  $f_+(e-2) = f_+(e+2) = 1/2$ . Therefore,

$$f_-(l) = \begin{cases} \frac{x}{2} & l = e \pm 3 \\ \frac{1-2x}{2} & l = e \pm 2 \\ \frac{x}{2} & l = e \pm 1 \\ 0 & \text{otherwise.} \end{cases} \quad (6.3.10)$$

From Equations (6.3.10) and (1.1.1), when  $f'_- \in \mathcal{J}_-^2$ , the entropy of  $f'_-$  is

$$H(f'_-) = -2x \log_2(x/2) - (1-2x) \log_2((1-2x)/2)$$

**Lemma 6.2.** *The entropy of any  $f_- \in \mathcal{J}_-^2$  is greater than the entropy of any  $f'_- \in \mathcal{J}_-^1$ .*

*Proof.* From Equations (6.3.6) and (6.3.11), if  $f_- \in \mathcal{J}_-^2$  and  $f'_- \in \mathcal{J}_-^1$ ,

$$\begin{aligned} H(f_-) - H(f'_-) &= -x \log_2(x/2) + x \log_2(x) \\ &= x(\log_2(x) - \log_2(x/2)) \\ &= x \log_2(2) \\ &> 0 \end{aligned} \quad (6.3.11)$$

The third equality holds because  $0 < x$ . □

Let us turn our attention to what happens when the creature chooses to move to certain locations. It is important to remember that when the creature decides where

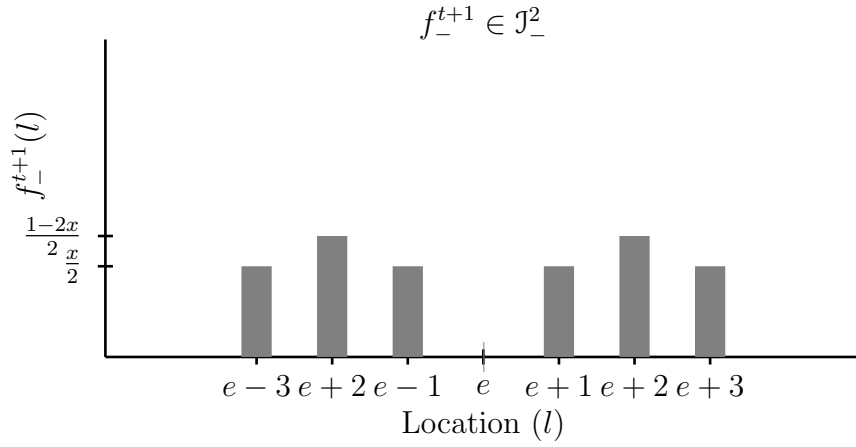


Figure 6.6: A typical probability mass function  $f_-$  in the information state  $\mathcal{J}_-^2$ . The creature is in the information state  $\mathcal{J}_-^2$  at time  $t$  if and only if  $F_+^t \in \mathcal{J}_-^2$ .

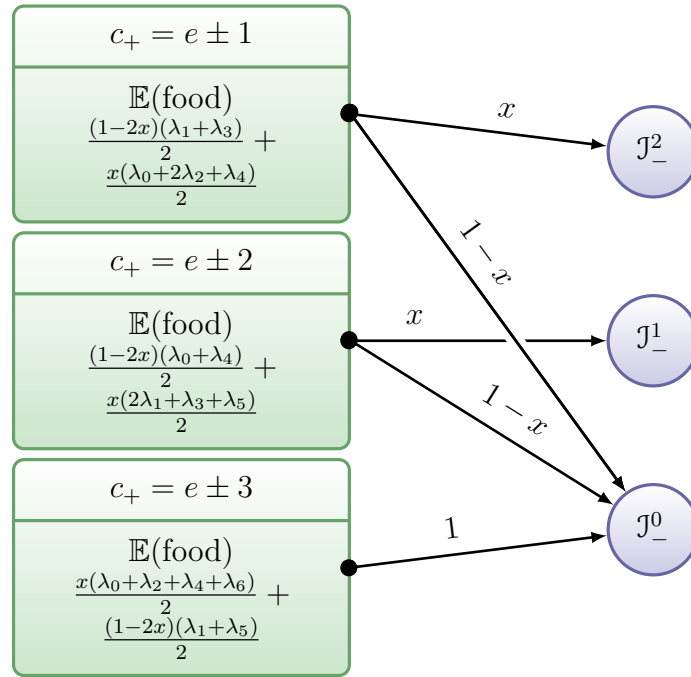


Figure 6.7: Transitions from  $\mathcal{J}_-^2$ . This figure states the expected food at each choice for the creature's next location ( $c_+$ ) when in the information state  $\mathcal{J}_-^1$ . It also depicts the transition probabilities to other information states based on  $c_+$ .

to move to it has access to the probability distribution of the injector's next location, namely  $f_-^{t+1}$ . When  $f_-^{t+1} \in \mathcal{J}_-^2$ , the creature has a set of expected results for moving to different locations on the ring. See Figures 6.6 and 6.7 for a visual representation.

1. Location  $C^{t+1} = c_+ \in \{e \pm 3\}$ .

Without loss of generality assume  $c_+ = e + 3$ .

The expected food for a creature that moves to  $c_+$  given  $f_-^t \in \mathcal{J}_-^0$  is the sum over  $l$  of the probability the injector is at  $l$  times the expected amount of food the creature will get when it is at  $c_+$  and the injector is at  $l$ .

$$\begin{aligned} \mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+-l}) \\ &= \frac{x(\lambda_0 + \lambda_2 + \lambda_4 + \lambda_6)}{2} + \frac{(1-2x)(\lambda_1 + \lambda_5)}{2} \end{aligned} \quad (6.3.12)$$

Using the same arguments present in Section 6.3.1, if the creature goes to  $e \pm 3$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1$$

2. Location  $C^{t+1} = c_+ \in \{e \pm 2\}$ .

Without loss of generality assume  $c_+ = e + 2$ .

$$\begin{aligned} \mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+-l}) \\ &= \frac{(1-2x)(\lambda_0 + \lambda_4)}{2} + \frac{x(2\lambda_1 + \lambda_3 + \lambda_5)}{2} \end{aligned} \quad (6.3.13)$$

Using the same arguments present in Section 6.3.1, if the creature goes to  $e \pm 2$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = x$$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1 - x$$

3. Location  $C^{t+1} = c_+ = e \pm 1$ .

$$\begin{aligned}
\mathbb{E}(\text{food}) &= \sum_l f_-^t(l) \mathbb{E}(K_{c_+ - l}) \\
&= \frac{(1 - 2x)(\lambda_1 + \lambda_3)}{2} + \frac{x(\lambda_0 + 2\lambda_2 + \lambda_4)}{2}
\end{aligned} \tag{6.3.14}$$

Using the same arguments present in Section 6.3.1, if the creature goes to  $e$

$$\begin{aligned}
\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) &= x \\
\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) &= 1 - x
\end{aligned}$$

### 6.3.4 The Information State $\mathcal{J}_-^{3+}$

The information state  $\mathcal{J}_-^{3+} := \{f_- \in \mathcal{P}(\mathbb{Z}_N) : f_- = f_+ * P_x \text{ with } f_+ \in \mathcal{J}_+^{3+}\}$ . Let us look at some of the properties of  $f_- \in \mathcal{J}_-^{3+}$ . There exists  $f_+ \in \mathcal{J}_+^{3+}$  such that  $f_- = f_+ * P_x$ . Therefore, there is a location  $e$  and distance  $2 < d \leq N/2$  such that  $f_+(e - d) = f_+(e + d) = 1/2$ . Therefore,

$$f_-(l) = \begin{cases} \frac{x}{2} & l = e \pm d \pm 1 \\ \frac{1-2x}{2} & l = e \pm d \\ 0 & \text{otherwise.} \end{cases} \tag{6.3.15}$$

From Equations (6.3.15) and (1.1.1), when  $f'_- \in \mathcal{J}_-^{3+}$ , the entropy of  $f'_-$  is

$$H(f'_- x) = -2x \log_2(x/2) - (1 - 2x) \log_2((1 - 2x)/2) \tag{6.3.16}$$

**Lemma 6.3.** *The entropy of any  $f_- \in \mathcal{J}_-^{3+}$  is equal to the entropy of any  $f'_- \in \mathcal{J}_-^2$ .*

*Proof.* To see this is true, compare Equations (6.3.16) and (6.3.11).  $\square$

The expected amount of food in this information state is dependent on  $d$ . Furthermore, none of the creature strategies implemented in this thesis spend more than one time step in this state. For these reasons the expected food is not explicitly stated. As a note to the reader the expected food at any location in this state is strictly less

than the expected food in the comparable location when in the state  $\mathcal{J}_-^2$ . See Figures 6.8 and 6.9 for a visual representation.

The only property of importance to further analysis is the transition probabilities.

1. Location  $C^{t+1} = c_+ \in \{e \pm d \pm 1\}$

If the creature goes to location  $c_+$ ,

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1$$

2. Location  $C^{t+1} = c_+ \in \{e \pm d\}$

If the creature goes to location  $c_+$ ,

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = x$$

$$\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1 - x$$

### 6.3.5 The Information State $\mathcal{J}_-^I$

The information state  $\mathcal{J}_-^I := \{f_- \in \mathcal{P}(\mathbb{Z}_N) : f_- = f_+ * P_x \text{ with } f_+ \in \mathcal{J}_+^I\}$ . Let us look at some of the properties of  $f_- \in \mathcal{J}_-^I$ . There exists  $f_+ \in \mathcal{J}_+^I$  such that  $f_- = f_+ * P_x$ . Therefore, there is a location  $e$  and distance  $0 < d \leq N/2$  such that  $f_+(e+d) = \beta > 0$ ,  $f_+(e-d) = 1 - \beta > 0$  and  $\beta \neq 1 - \beta$ .

1.  $d = 1$

When the creature knows that the injector is almost surely 1 step away from location  $e$ ,  $f_- \in \mathcal{J}_-^I$  is of the form

$$f_-(l) = \begin{cases} \beta x & l = e - 2 \\ \beta(1 - 2x) & l = e - 1 \\ x & l = e \\ (1 - \beta)(1 - 2x) & l = e + 1 \\ (1 - \beta)x & l = e + 2 \\ 0 & \text{otherwise.} \end{cases} \quad (6.3.17)$$

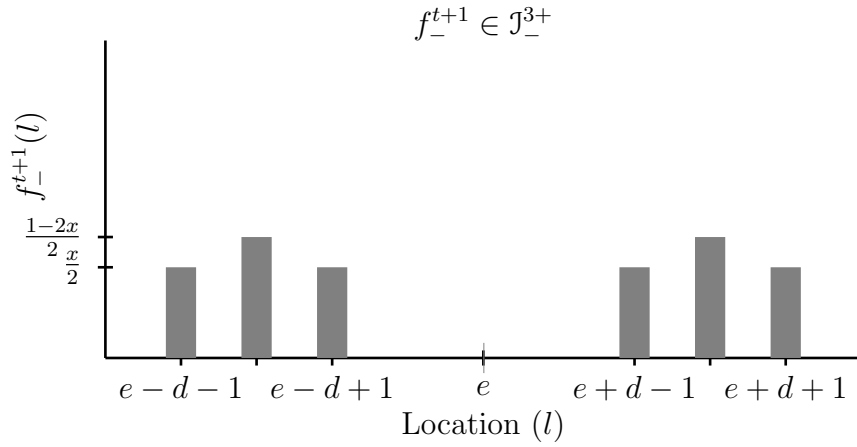


Figure 6.8: A typical probability mass function  $f_-$  in the information state  $\mathcal{J}_-^{3+}$ . The creature is in the information state  $\mathcal{J}_-^{3+}$  at time  $t$  if and only if  $F_+^t \in \mathcal{J}_-^{3+}$ .

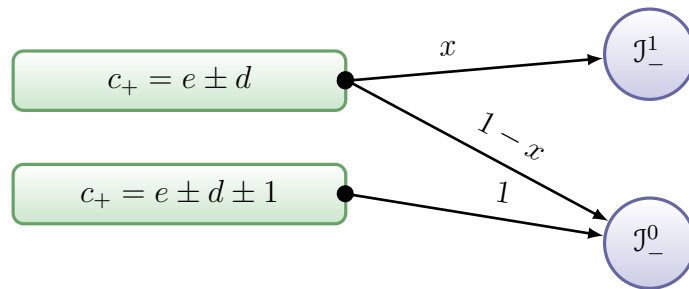


Figure 6.9: Transitions from  $\mathcal{J}_-^{3+}$ . The expected food in this state is a function of  $d$ ; the creature strategies do not spend more than one turn in this state so the effect on the expected food over time is negligible. It does depict the transition probabilities to other information states based on  $c_+$ .

2.  $d > 1$

When the creature knows that the injector is almost surely 2 or more steps away from location  $e$ ,  $f_- \in \mathcal{J}_-^I$  is of the form

$$f_-(l) = \begin{cases} \beta x & l = e - d \pm 1 \\ \beta(1 - 2x) & l = e - d \\ (1 - \beta)(1 - 2x) & l = e + d \\ (1 - \beta)x & l = e + d \pm 1 \\ 0 & \text{otherwise.} \end{cases} \quad (6.3.18)$$

In the other information states we were able to claim universal properties of all probability mass functions within those information states but that is not case for  $\mathcal{J}_-^I$ . The entropy and the expected amount of food at any given location will be dependent upon  $\beta$ . In the

**Lemma 6.4.** *The entropy of any  $f_- \in \mathcal{J}_-^I$  is greater than the entropy of any  $f'_- \in \mathcal{J}_-^0$ .*

*Proof.* The entropy of  $f_- \in \mathcal{J}_-^I$  with  $d = 1$  is

$$\begin{aligned} H(f_-) &= -\beta x \log_2(\beta x) - \beta(1 - 2x) \log_2(\beta(1 - 2x)) \\ &\quad - x \log_2(x) - (1 - \beta)x \log_2((1 - \beta)x) \\ &\quad - (1 - \beta)(1 - 2x) \log_2((1 - \beta)(1 - 2x)) \end{aligned}$$

Assume (without loss of generality) that  $\beta > 1 - \beta$ . Consider the term

$$\begin{aligned} & -\beta x \log_2(\beta x) - (1 - \beta)x \log_2((1 - \beta)x) \\ &= -x \log_2((1 - \beta)x) - \beta x \log_2(\beta x) + \beta x \log_2((1 - \beta)x) \\ &= -x \log_2((1 - \beta)x) + \beta x \log_2\left(\frac{1 - \beta}{\beta}\right) \\ &> -x \log_2((1 - \beta)x) \end{aligned} \quad (6.3.19)$$

Similarly,

$$\begin{aligned} & -\beta(1 - 2x) \log_2(\beta(1 - 2x)) - (1 - \beta)(1 - 2x) \log_2((1 - \beta)(1 - 2x)) \\ &> -(1 - 2x) \log_2(1 - 2x) \end{aligned} \quad (6.3.20)$$



Therefore,

$$\begin{aligned} H(f_-) &> -x \log_2(x) - x \log_2(x) - (1 - 2x) \log_2(1 - 2x) \\ &= H(f'_-) \quad \forall f'_- \in \mathcal{J}_-^0 \end{aligned}$$

The entropy of  $f_-^* \in \mathcal{J}_-^I$  with  $d > 1$  is

$$\begin{aligned} H(f_-) &= -2\beta x \log_2(\beta x) - \beta(1 - 2x) \log_2(\beta(1 - 2x)) \\ &\quad - 2(1 - \beta)x \log_2((1 - \beta)x) \\ &\quad - (1 - \beta)(1 - 2x) \log_2((1 - \beta)(1 - 2x)) \end{aligned}$$

The difference between the entropy of  $f_-^* \in \mathcal{J}_-^I$  with  $d > 1$  and  $f_- \in \mathcal{J}_-^I$  with  $d = 1$  is

$$\begin{aligned} H(f_-^*) - H(f_-) &= -\beta x \log_2(\beta x) - (1 - \beta)x \log_2((1 - \beta)x) + x \log_2(x) \\ &> -x \log_2(x) + x \log_2(x) \\ &= 0 \end{aligned}$$

The first inequality holds from Equation (6.3.19).

Therefore we have shown for all  $f_-^* \in \mathcal{J}_-^I$  with  $d > 1$ , for all  $f_- \in \mathcal{J}_-^I$  with  $d = 1$  and for all  $f'_- \in \mathcal{J}_-^0$

$$H(f_-^*) > H(f_-) > H(f'_-) \quad \square$$

### 6.3.6 Entropy Minimal State

An information state is defined to be entropy minimal when the entropy of every probability mass function in that state is strictly less than a probability mass function in any other information state.

**Proposition 6.5.** *The information state  $\mathcal{J}_-^0$  is entropy minimal.*

*Proof.* Let  $f'_-$  in  $\mathcal{J}_0$ . First let us consider  $f_- \in \mathcal{J}_-^*$ . The entropy of  $f_-^0$  is  $\log_2 N$ . For  $N > 3$ ,  $\log_2 N > -2x \log_2(x) - (1 - 2x) \log_2(1 - 2x) = H(f'_x)$ . To see this, simply note that  $x = 1/3$  maximizes  $H(f'_x)$  and that  $\log_2$  is a monotonically increasing function. Therefore for all  $N > 3$ ,  $\log_2 N > \log_2 3 \geq H(f'_x)$ .

For all  $f_- \in \mathcal{J}_-^1 \cup \mathcal{J}_-^I$  the entropy of  $f_-$  is strictly greater than the entropy of  $f'_-$ .  
This follows from Lemmas 6.1 and 6.4.

For all  $f_- \in \mathcal{J}_-^2 \cup \mathcal{J}_-^{3+}$  the entropy of  $f_-$  is strictly greater than the entropy of  $f'_-$ .  
This follows from Lemmas 6.1, 6.2 and 6.3. □

## 7 Creature Strategies

A creature's strategy defines how it moves throughout the world given its history. The strategy can be a specific set of rules (e.g. when you have some history  $h^t$ , move to this location) or it can be a set of guiding principles (e.g. optimize your average information about the injector's location). When stating a set of guiding principles, it can be challenging to find the specific set of rules that carries out the principle. The work in Sections 5 and 6 has reduced every supported  $h^t$  to a probability mass function  $f_-^{t+1}$  and then classified every supported  $f_-^{t+1}$  into an information state. Now a strategy can be implemented simply by saying where to move given  $f_-^{t+1}$  for each of the information states.

This section will make heavy use of the primary theorem of Markov Processes; namely, if there exist a finite number of states with transition probabilities between the states, there is a nonnegative eigenvector with an eigenvalue of 1 and whose components sum to 1 that represents the probability of being in each state when the system is in equilibrium. Each strategy considered leads to a Markov chain on a finite subset of the information states defined in Section 6. As seen in Figures 6.3, 6.5, 6.7 and 6.9, if we create a movement rule for each information state, we will have a set of transition probabilities from each state to the other states.<sup>7</sup> Therefore, we will be able to calculate the probability that  $F_-^{t+1}$  is in each of those information states at steady state. Furthermore, we have calculated the average amount of food yielded per timestep by each strategy in the long run (*i.e.* at the equilibrium distribution for the corresponding Markov chain) as well as the average entropy of the probability mass functions representing the precision with which the creature knows the injector's location at each time step.

---

<sup>7</sup>We do not include a movement rule  $\mathcal{J}_-^I$  because it is unsupported for strategies in this thesis.

## 7.1 Information

Although the definition of the mutual information between the injector's next location and its history was presented in Section 1.1, it is not until now that all the tools necessary to calculate the information have been presented. From Equation (1.1.3) and for a given time  $t$ , the mutual information between the injector's next location  $S^{t+1}$  and its history is

$$H(S^{t+1}; \mathcal{H}^t) = H(S^{t+1}) - H(S^{t+1}|\mathcal{H}^t)$$

As shown in Section 1.1 the entropy of  $S^{t+1}$  for all  $t$  is  $\log(N)$ . Looking at the second term,

$$H(S^{t+1}|\mathcal{H}^t) = \mathbb{E}_{h^t} H(S^{t+1}|h^t) \quad \text{See Equation (1.1.2)} \quad (7.1.1)$$

From the work in Section 4 this can be reduced to

$$\mathbb{E}_{h^t} H(S^{t+1}|h^t) = \mathbb{E}_{f_-^{t+1}} H(f_-^{t+1}) \quad (7.1.2)$$

$$= H(F_-^{t+1}) \quad (7.1.3)$$

When the creature is in the steady state defined by the Markov Process,

$$\begin{aligned} H(S; \mathcal{H}) &= \mathbb{E}_{J_-} (H(S^{t+1}) - H(F_-^{t+1})) \\ &= \log(N) - \mathbb{E}_{J_-} (H(F_-^{t+1})) \end{aligned} \quad (7.1.4)$$

We introduce the following notation. The probability that  $f_- \in \mathcal{J}_-^a$  given the creature is following some strategy will be denoted by  $\mathbb{P}_{strat}(\mathcal{J}_-^a)$ . The entropy of an information state  $\mathcal{J}_-^a$  is the entropy of each  $f_- \in \mathcal{J}_-^a$  and will be denoted by  $H(\mathcal{J}_-^a)$ . If  $a \in \{0, 1, 2\}$  then  $H(\mathcal{J}_-^a)$  is well defined.

If a strategy is a Markov Process on the states  $\mathcal{J}_-^0$ ,  $\mathcal{J}_-^1$  and  $\mathcal{J}_-^2$ , then

$$H(F_-^t) = \mathbb{P}_{strat}(\mathcal{J}_-^0)H(\mathcal{J}_-^0) + \mathbb{P}_{strat}(\mathcal{J}_-^1)H(\mathcal{J}_-^1) + \mathbb{P}_{strat}(\mathcal{J}_-^2)H(\mathcal{J}_-^2) \quad (7.1.5)$$

The probability that  $f_-$  is in each information state for a given strategy is determined by the Markov transition matrix. The entropy of each information state was calculated in Section 6.

## 7.2 Information Theory Creature

The guiding principle for the Information Theory Creature (ITC) is to maximize its information about the injector's next location (over all time) given the creature knows its full history. In order to maximize the information the creature must choose a strategy that minimizes the entropy of  $F_-^{t+1}$  over time (See Equation 7.1.5). From Proposition 6.5 the information state  $\mathcal{J}_-^0$  is entropy minimal. Therefore, if there is a set of movement rules where  $f_-^{t+1}$  is almost surely in the information state  $\mathcal{J}_-^0$ , for all  $t > 0$ , then this set of rules describes an information optimal strategy.

Consider the following set of rules:

1.  $f_-^{t+1} \in \mathcal{J}_-^0$ 
  - (a) Rule: The creature moves to  $e \pm 1$ .
  - (b) Transition:  $f_-^{t+2}$  is almost surely in  $\mathcal{J}_-^0$ .
  - (c) Expected Food:  $x\lambda_0 + (1 - 2x)\lambda_1 + x\lambda_2$ .
2.  $f_-^{t+1} \in \mathcal{J}_-^1$ 
  - (a) Rule: The creature moves to  $e \pm 2$ .
  - (b) Transition:  $f_-^{t+2}$  is almost surely in  $\mathcal{J}_-^0$ .
3.  $f_-^{t+1} \in \mathcal{J}_-^2$ 
  - (a) Rule: The creature moves to  $e \pm 3$ .
  - (b) Transition:  $f_-^{t+2}$  is almost surely in  $\mathcal{J}_-^0$ .
4.  $f_-^{t+1} \in \mathcal{J}_-^3$ 
  - (a) Rule: The creature moves to  $e \pm d \pm 1$ .
  - (b) Transition:  $f_-^{t+2}$  is almost surely in  $\mathcal{J}_-^0$ .

Looking at Figure 7.1, after two time steps the creature is almost surely in and almost surely stays in the information state  $\mathcal{J}_-^0$ . Therefore, this strategy is information optimal. The expected food for this strategy is  $x\lambda_0 + (1 - 2x)\lambda_1 + x\lambda_2$ .

### 7.3 Maximum Likelihood Creature

This Maximum Likelihood Creature (MLC) goes to the location that the injector is most likely to be.<sup>8</sup> This is plausibly a good strategy because there is almost surely more food at the injector's location than any where else.

When we restrict  $P_x$  such that  $1/4 < x < 1/3$  the following set of rules constitutes a Maximum Likelihood strategy. Consider the following set of rules:

1.  $f_-^{t+1} \in \mathcal{J}_-^0$ 
  - (a) Rule: The creature moves to  $e$ .
  - (b) Transition:  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1 - 2x$  and  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = 2x$ .
  - (c) Expected Food:  $(1 - 2x)\lambda_0 + 2x\lambda_1$
2.  $f_-^{t+1} \in \mathcal{J}_-^1$ 
  - (a) Rule: The creature moves to  $e$ .
  - (b) Transition:  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = x$ ,  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = 1 - 2x$  and  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^2) = x$ .
  - (c) Expected Food:  $x\lambda_0 + (1 - 2x)\lambda_1 + x\lambda_2$ .
3.  $f_-^{t+1} \in \mathcal{J}_-^2$ 
  - (a) Rule: The creature moves to  $e \pm 2$ .
  - (b) Transition:  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1 - x$  and  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = x$ .
  - (c) Expected Food:  $\frac{(1-2x)(\lambda_0+\lambda_4)}{2} + \frac{x(2\lambda_1+\lambda_3+\lambda_5)}{2}$ .

---

<sup>8</sup>In case there are  $n > 1$  such locations, it chooses between them with probability  $1/n$  each. Note the transitions specified for state  $\mathcal{J}_-^{3+}$  are an exception to the rule of moving to a maximum of  $F_-^t$ . This exception is required to obtain a closed Markov chain on three states.

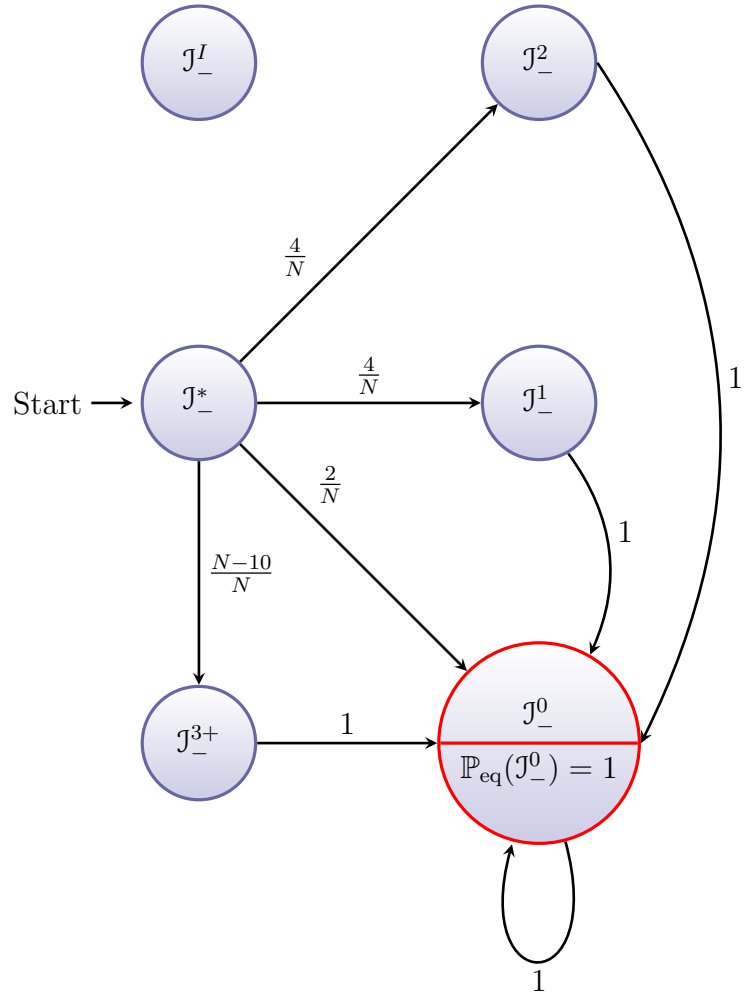


Figure 7.1: Information Theory Creature State Transitions. This figure depicts the transition probabilities for the Information Theory strategy. After 2 time steps the creature is almost most surely in  $\mathcal{J}_-^0$  and will almost surely stay there. The starting distribution  $\mathcal{J}_-^*$  corresponds to  $1/N$  for each location, *i.e.*  $f_-^0(s) = 1/N$  for all  $s$ .

4.  $f_-^{t+1} \in \mathcal{J}_-^{3+}$

(a) Rule: The creature moves to  $e \pm d$ .

(b) Transition:  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1 - x$  and  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = x$ .

Looking at Figure 7.2, after two time steps the creature is almost surely in  $\mathcal{J}_-^0$ ,  $\mathcal{J}_-^1$  or  $\mathcal{J}_-^2$  and will almost surely stay in one of those three states. Therefore, this strategy forms a Markov process on those states and the transitions can be expressed by the following matrix.

$$T_{MLC} = \begin{bmatrix} 1 - 2x & x & 1 - x \\ 2x & 1 - 2x & x \\ 0 & x & 0 \end{bmatrix} \quad (7.3.1)$$

Where the  $lm$  entry is the probability that creature in  $\mathcal{J}_-^{m-1}$  goes to  $\mathcal{J}_-^{l-1}$ . By the Perron-Frobenius theorem, there exists a nonnegative vector

$$\hat{\mathbb{P}}_{MLC} = \begin{bmatrix} \mathbb{P}_{MLC}(\mathcal{J}_-^0) \\ \mathbb{P}_{MLC}(\mathcal{J}_-^1) \\ \mathbb{P}_{MLC}(\mathcal{J}_-^2) \end{bmatrix}^*$$

such that  $\hat{\mathbb{P}}_{MLC} \cdot T_{MLC} = \hat{\mathbb{P}}_{MLC}$ . The vector

$$\hat{\mathbb{P}}_{MLC} = \begin{bmatrix} \mathbb{P}_{MLC}(\mathcal{J}_-^0) \\ \mathbb{P}_{MLC}(\mathcal{J}_-^1) \\ \mathbb{P}_{MLC}(\mathcal{J}_-^2) \end{bmatrix}^* = \begin{bmatrix} \frac{2-x}{4+x} \\ \frac{2}{4+x} \\ \frac{2x}{4+x} \end{bmatrix}^* \quad (7.3.2)$$

is the unique vector satisfying that equation. From Proposition 6.5, Equation 7.1.5 and the fact that the Maximum Likelihood Creature has a non-zero probability of being in  $\mathcal{J}_-^1$  and  $\mathcal{J}_-^2$ , this strategy has strictly worse information about the injector's next location than the Information Theory strategy.

## 7.4 Modified Maximum Likelihood Creature

The modified Maximum Likelihood Creature (mMLC) is a hybridization of the information theory strategy and the maximum likelihood strategy. When the creature is



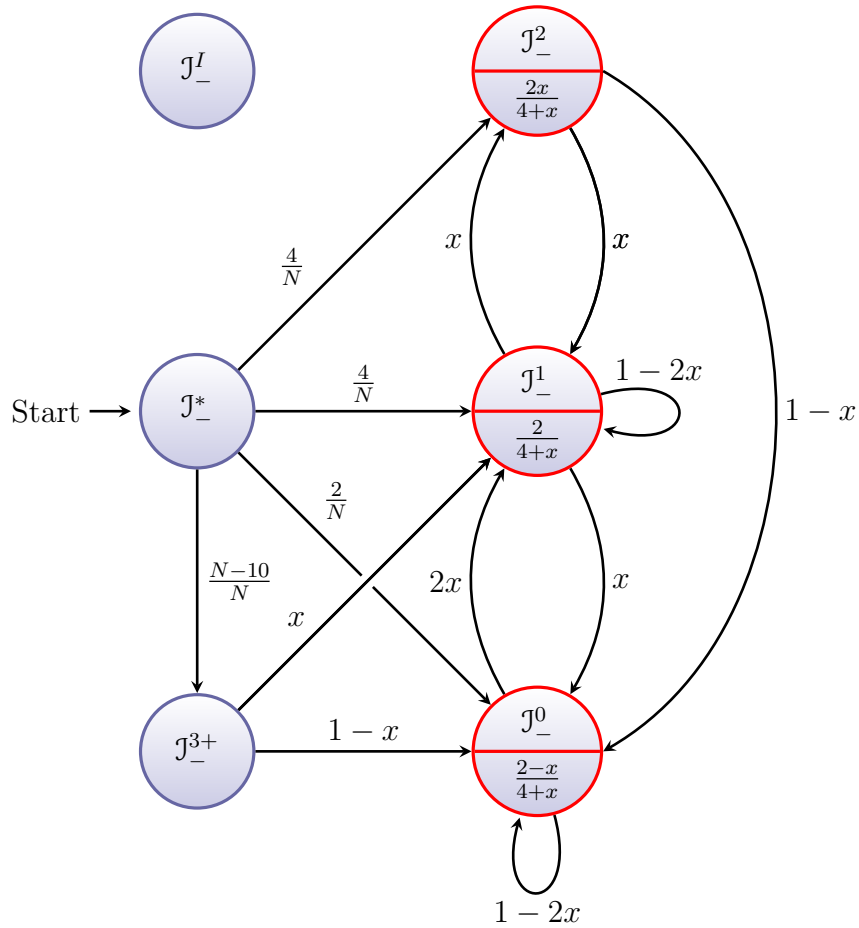


Figure 7.2: Maximum Likelihood Creature State Transitions. This figure depicts the transition probabilities for the Maximum Likelihood strategy. After 2 time steps the creature is almost most surely in  $\mathcal{J}^0_-$ ,  $\mathcal{J}^1_-$  or  $\mathcal{J}^2_-$ , and will almost surely stay within those three states (though it clearly may move between them). The steady state proportion of time spent in each of these states is listed below the name.

in  $\mathcal{J}_-^0$  it goes wherever the injector is most likely to be; if it is not in  $\mathcal{J}_-^0$  it tries to get there. When we restrict  $P_x$  such that  $x < 1/3$  the following set of rules constitutes a modified Maximum Likelihood strategy. Consider the following set of rules:

1.  $f_-^{t+1} \in \mathcal{J}_-^0$ 
  - (a) Rule: The creature moves to  $e$ .
  - (b) Transition:  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1 - 2x$  and  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^1) = 2x$ .
  - (c) Expected Food:  $(1 - 2x)\lambda_0 + 2x\lambda_1$
2.  $f_-^{t+1} \in \mathcal{J}_-^1$ 
  - (a) Rule: The creature moves to  $e \pm 2$ .
  - (b) Transition:  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1$ .
  - (c) Expected Food:  $\frac{x(\lambda_0 + 2\lambda_2 + \lambda_4)}{2} + \frac{(1-2x)(\lambda_1 + \lambda_3)}{2}$ .
3.  $f_-^{t+1} \in \mathcal{J}_-^2$ 
  - (a) Rule: The creature moves to  $e \pm 3$ .
  - (b) Transition:  $\mathbb{P}(f_-^{t+2} \in \mathcal{J}_-^0) = 1$ .
4.  $f_-^{t+1} \in \mathcal{J}_-^{3+}$ 
  - (a) Rule: The creature moves to  $e \pm d \pm 1$ .
  - (b) Transition:  $f_-^{t+2}$  is almost surely in  $\mathcal{J}_-^0$ .

Looking at Figure 7.3, after two time steps the creature is almost surely in  $\mathcal{J}_-^0$  or  $\mathcal{J}_-^1$  and will almost surely stay in one of those two states. Therefore, this strategy almost surely converges within two time steps to a Markov process on those states and the transitions can be expressed by the following matrix.

$$T_{mMLC} = \begin{bmatrix} 1 - 2x & 1 \\ 2x & 0 \end{bmatrix} \quad (7.4.1)$$

Where the  $lm$  entry is the probability that creature in  $\mathcal{J}_-^{m-1}$  goes to  $\mathcal{J}_-^{l-1}$ . Because this is a Markov transition matrix there exists a vector

$$\hat{\mathbb{P}}_{mMLC} = \begin{bmatrix} \mathbb{P}_{mMLC}(\mathcal{J}_-^0) \\ \mathbb{P}_{mMLC}(\mathcal{J}_-^1) \end{bmatrix}^*$$

such that  $\hat{\mathbb{P}}_{mMLC} \cdot T_{mMLC} = \hat{\mathbb{P}}_{mMLC}$ . The vector

$$\hat{\mathbb{P}}_{mMLC} = \begin{bmatrix} \mathbb{P}_{mMLC}(\mathcal{J}_-^0) \\ \mathbb{P}_{mMLC}(\mathcal{J}_-^1) \end{bmatrix}^* = \begin{bmatrix} \frac{1}{1+2x} \\ \frac{2x}{1+2x} \end{bmatrix}^* \quad (7.4.2)$$

is the unique vector satisfying that equation. This strategy still has less information about the injector's next location than the Information Theory strategy but more information than Maximum Likelihood strategy (See Figure 7.4).

## 7.5 Results

This section analyzes the average amount of information and food each type of creature is expected to have in the distance-certain model with the standard parameters listed in Table 2.1. As seen in Figure 7.4, the information theory creature has strictly better information than the modified maximum likelihood creature which has strictly better information than the maximum likelihood creature. This is as expected because we proved the information theory creature has optimal information and the modified maximum likelihood creature is a hybrid of the other two strategies. As the injector becomes more mobile (i.e.  $x$  goes to  $1/3$ ) the amount of information about the injector's location decreases for all three strategies. The injector's next location is most uncertain when  $x = 1/3$ . Again, this is what one would intuitively expect about the expected information for these strategies.

With regards to the expected food, none of the strategies is strictly better than the other (See Figure 7.5). In fact, there are ranges of  $P_x$  in which each strategy is strictly better than the others. Furthermore, as the injector becomes more mobile and the creature has less information about the injector's next location, the modified

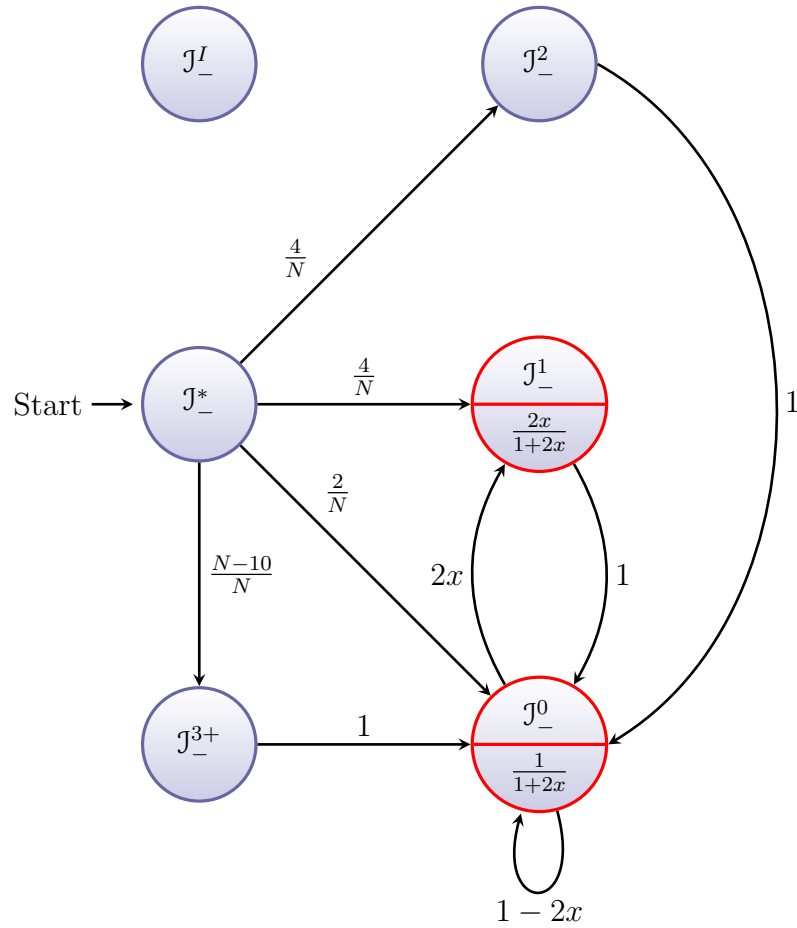


Figure 7.3: Modified Maximum Likelihood Creature State Transitions. This figure depicts the transition probabilities for the modified Maximum Likelihood strategy. After 2 time steps the creature is almost most surely in  $J^0_-$  or  $J^1_-$ , and will almost surely stay within those two states (though it clearly may move between them). The steady state proportion of time spent in each of these states is listed below the name.

maximum likelihood creature and the maximum likelihood creature do worse in terms of food gathering, but the information theory creature does better.

One might hypothesize that information becomes more useful as the environment becomes more uncertain, but that would fail to explain the expected food gathered by the modified maximum likelihood creature. Although it has the second best information, it is worst of the three when the environment is more uncertain and the best when it is more certain. The fact that the modified maximum likelihood creature is not an average of the other two strategies in terms of food gathered indicates there isn't a simple trade off between food and information. This is certainly surprising if one accepts information theory as an explanatory principle.

This does not mean that information is unimportant to the creature's survival. If the creature had no ability to predict the injector's location, its movement would have to be independent of the injector's location. In that case the creature would see on average  $\frac{\gamma}{N\alpha} \approx 0.138\gamma$  molecules (using standard parameters and  $\alpha = 0.4$ )<sup>9</sup>. This is substantially less than any of the other strategies.

Since information is not a good way to predict a strategy's success, it is worth looking to the Markov process for an explanation. There are two components for determining a strategy's average food: the expected food for each movement rule in each state and the relative amount of time in each state.

For information state  $\mathcal{J}_-$ , when  $x < 1/3$  the expected food at the center location is always better than the expected food at the outer locations. This means that the information theory creature does worse in state  $\mathcal{J}_-$  than the other two strategies (See Figure 6.3). That being said, the difference between the two locations decreases as  $x$  increases from zero to one third. The reason that the maximum likelihood and modified maximum likelihood strategies are not always better than the information theory creature is they don't spend all their time in  $\mathcal{J}_-$ . In fact, as  $x$  goes from 1/4 to

---

<sup>9</sup>There are on average  $\gamma/\alpha$  molecules in the world (See Section 3.4) that are equally likely to be at any of  $N$  locations.

1/3 the maximum likelihood and modified maximum likelihood creatures spend less time in  $\mathcal{J}_-^0$  (See Figures 7.2 and 7.3). The reason the ITC does better as  $x$  gets bigger is that it spends more and more time on top of the injector as  $x$  gets larger; it spends a fraction of time on the injector equal to  $x$ .

For information state  $\mathcal{J}_-^1$ , the maximum likelihood creature does just as well as the ITC in terms of food while the mMLC does significantly worse (See Figures 6.3 and 6.5). Moreover, the mMLC spends a greater percentage of its time in  $\mathcal{J}_-^1$  as  $x$  goes from 0 to 1/3 (See Figure 7.3).

For information state  $\mathcal{J}_-^2$ , the maximum likelihood creature does worse than the ITC in  $\mathcal{J}_-^0$  and as  $x$  goes from 1/4 to 1/3 (See Figures 6.3 and 6.7). In addition it spends more time in this state (See Figure 7.2).

The reason for the food difference between the maximum likelihood and modified maximum likelihood strategies is that as  $x$  goes to 1/4 the modified maximum likelihood creature spends significantly more of its time in  $\mathcal{J}_-^0$  than the maximum likelihood creature (See Figures 7.2 and 7.3).

We conjecture that for the model discussed with finite values of the injection rate  $\gamma$ , there is a sufficiently large value of  $\gamma$  such that the performance of the different strategies in that case is consistent with the performance of the same strategies in the distance-certain case presented here.

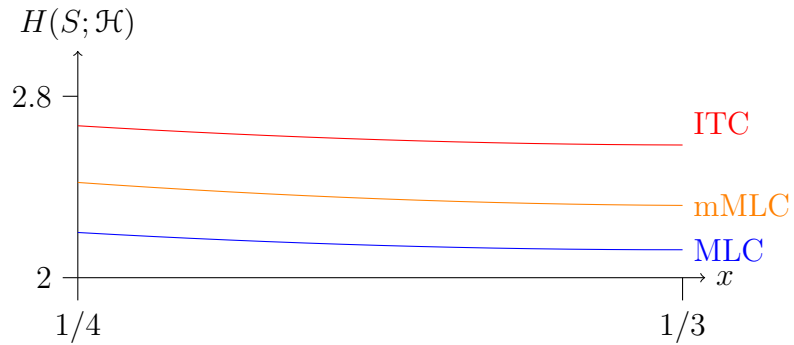


Figure 7.4: This graph depicts the expected information in the long term for each strategy given  $P_x$ .

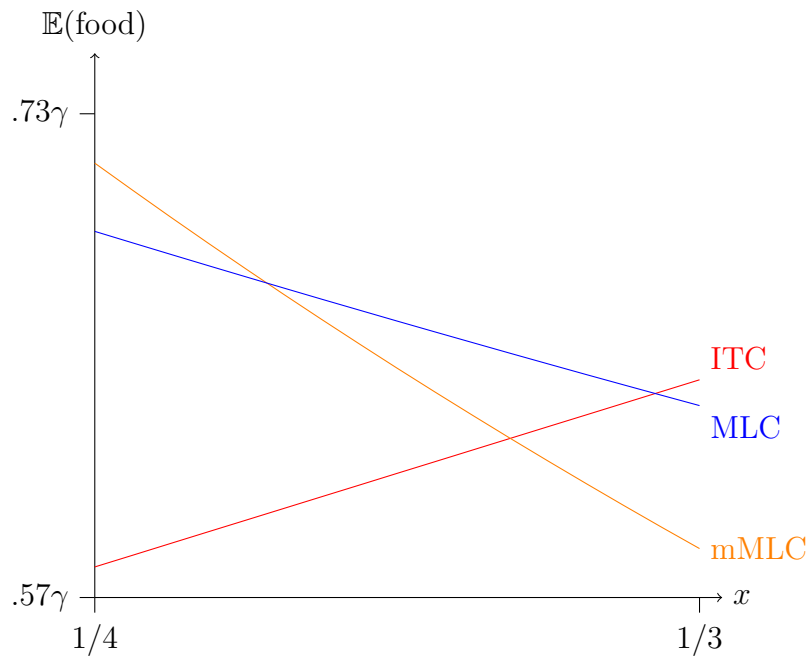


Figure 7.5: This graph depicts the average amount of food seen by each strategy at each observation. This is a function of  $P_x$  and  $\gamma$  (as listed explicitly) as well as  $N$ ,  $Q_y$  and  $\alpha$ . We use the standard parameter values listed in Table 2.1 with  $\alpha = .4$ .

## 8 Discussion

### 8.1 Alternative Food Distributions

This model was built on a particular choice for molecule injection, movement and decay. It followed that the food distributed about the injector with an (almost) exponential decay in the average food at each location. Section 4 only requires that the food at each distance from the injector have some known distribution. If these distributions have a parallel to the results in Section 5 where high injection rates mean the creature has a good estimate of the injector's location, then one could try and apply the methods in Sections 6 and 7.

### 8.2 The Injector's Movement Algorithm

In this model we only consider the case where the injector moves, the creature moves and makes an observation and then the injector moves again. However, there are other reasonable scenarios that can be described by the mathematics of this model.

First, let us consider a creature that can move to multiple locations before the injector can move. This is analogous to a creature feeding in tidal pools that are fixed between the tides. We will restrict consideration to a fixed number of observations per movement of the injector. In this case, an update of  $F_-^t$  to  $F_-^{t+1}$  using the  $t^{th}$  observation uses  $P_0$  when the injector could not move and  $P$  when the injector can.

Second, let us consider a creature that forages for some of the time and then sleeps for the night or engages in some other prolonged activity. While the creature is awake and foraging an update of  $F_-^t$  to  $F_-^{t+1}$  uses  $P$  but an update of  $F_-^t$  to  $F_-^{t+1}$  from dusk ( $t$ ) to dawn ( $t + 1$ ) uses  $P * P * \dots * P$ .  $P * P * \dots * P$  is  $P$  convolved with itself an appropriate number of iterations to cover the time from last foraging to the first observation of the next day/time/cycle.

Finally, let us consider a stochastic  $P$ . As described in this model,  $P$  is the same for all injector movements and is independent of all events in the world. This is



not entirely necessary. The creature must know how  $P$  is determined but that could include stochastic factors. For example, the injector could be more likely to move if it did not move last time and vice versa. Or the injector could try and move away from the creature.

### 8.3 Survival Time

The average survival time for each strategy is certainly a meaningful way to compare how successful the strategies are; that being said, surviving is not the same as thriving and the average food is still an important measure of the strategies' rankings. In order to gain an analytic result for survival time one could look at the Markov process on  $\mathcal{J}_-$  as driving a Markov process on the integers (representing food). Each time step would consume a number of molecules and the creature would gain an amount of food proportional to the number of molecules at its location. Mortality would correspond to reaching a lower bound (say, zero) for the creature's net food reserves. In this case the survival time would be formulated as a first passage time problem, *i.e.* the time for the expanded Markov chain to reach an absorbing boundary. Application of classical Markov chain techniques to this problem is a topic for future work. In addition, computer simulations of the system may prove indispensable for estimating the expected survival times for particular parameter values.

## 9 Conclusion

In this simple example, the set of information optimal strategies and the set of food optimal strategies can be mutually exclusive under mild assumptions. This conclusion is presented in Figures 7.5 and 7.4. There are injector movement algorithms (quantitatively but not qualitatively different) where the information optimal strategy is clearly not food optimal. Furthermore, there is no consistent relationship between information and expected food.

Although the techniques presented in this thesis only allow the discussion about the average results over the long term, this discussion is still useful for three reasons. Firstly, many creatures do live for long periods of time and so the results are directly applicable to them. Secondly, long term refers to a large number of time steps rather a long time. For example if the time steps were 1 second, in the course of an hour the average behavior would be representative. Thirdly, the argument that information optimal behavior is utility optimal behavior arises from the idea that “information is good” and therefore should be optimized. Whether we are discussing long term averages or average survival times the form of the argument is the same. Future work will use simulations to explore the average survival time for each strategy.

To make analysis tractable we assumed *in effect* that the injection rate went to infinity. This does not mean that there is a large amount nutrition in the world, just a large number of food particles. For example, as we increase the injection rate we could decrease the amount of energy obtained per molecule, while keeping the net available nutrition constant. This situation would mean the creatures are still trying to obtain a finite amount of energy or nutrition; therefore, strategies can be meaningfully compared by the average amount of food they garner even with high injection rates. Future work will use simulations to discover how successful the strategies are when the injection rate is finite.

## References

- [1] C. Adami. Introduction to artificial life. *IEEE Transactions on Evolutionary Computation*, 2(4):168–170, 1998.
- [2] A J Bell and T J Sejnowski. The "independent components" of natural scenes are edge filters. *Vision Res*, 37(23):3327–3338, Dec 1997.
- [3] M Dewese. Optimization principles for the neural code. *Network*, 7(2):325–331, May 1996.
- [4] R.A. Gatenby and B.R. Frieden. Application of information theory and extreme physical information to carcinogenesis, 2002.
- [5] RA Gatenby and BR Frieden. Information theory in living systems, methods, applications, and challenges. *Bull Math Biol*, 69(2):635–657, Feb 2007.
- [6] W. Huang, C. Ofria, and E. Torng. Measuring biological complexity in digital organisms. In *Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems*, page 315. Mit Press, 2004.
- [7] G. Jumarie. *Subjectivity, information, systems: an introduction to a theory of relativistic cybernetics*. Gordon & Breach Science Pub, 1986.
- [8] Y. Karklin and M.S. Lewicki. Emergence of complex cell properties by learning to generalize in natural scenes. *Nature*, November 2008.
- [9] J.F.C. Kingman. *Poisson Processes*. Oxford University Press, 1993.
- [10] C.E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3):379–423, 1948.
- [11] C.E. Shannon. A Mathematical Theory of Communication. *The Bell System Technical Journal*, 27(3):623–656, 1948.

- [12] E.C. Smith and M.S. Lewicki. Efficient Auditory Coding. *Nature*, 439(7079):978–982, 2006.
- [13] D R Tailor, L H Finkel, and G Buchsbaum. Color-opponent receptive fields derived from independent component analysis of natural images. *Vision Res*, 40(19):2671–2676, 2000.
- [14] O. Taussky. A recurring theorem on determinants. *American Mathematical Monthly*, 56(4):672–676, 1949.