Bayesian Framework for the Rendezvous Problem on the Line

Jeff Braun and John C. Wierman Johns Hopkins University

Abstract

A famous open problem in the field of rendezvous search is to ascertain the rendezvous value of the symmetric rendezvous problem on the line wherein both agents begin two units apart. We provide a new, Bayesian framework to both create new strategies for the agents to follow and to provide a new interpretation of previously posited strategies. Additionally, we have developed a method that modifies any strategy, even those with potentially infinite expected meeting time, into a new strategy that is guaranteed to have a finite expected meeting time. This process, combined with using our Bayesian framework to create new strategies, yields an upper bound that is within one percent of the current best upper bound for the symmetric rendezvous value.

1 The Problem

On an infinite line, two agents are placed two units apart. Neither agent has any knowledge as to which direction the other agent is in relative from their starting position, nor is there a universal notion of direction. Each agent has the ability to move at a maximum speed of one in each time unit. If both agents occupy the same space on the number line, we say that they have met or rendezvous has occurred. If each agent must adopt the same strategy (that is, they can not meet beforehand to take on different roles), which strategy will guarantee the smallest expected meeting time? Furthermore, what is that value, also known as the symmetric rendezvous value?

2 Current State of Research

The symmetric rendezvous search problem on the line has been an open problem for the past 25 years. Much progress has been made on tightening the upper and lower bounds on the smallest expected meeting time (denoted R^s), coming up with strategies that achieve surprisingly small expected meeting times, and narrowing the conditions which an optimal strategy must meet.

Alpern [1], who originally postulated the problem, derived an upper bound of 5.0. His strategy was for each agent to randomly decide which direction would be "forward", and then to move 1 unit forward and 2 units backwards in succession. If rendezvous has still not occurred, the entire process is repeated, including randomly selecting which direction is forward.

Han et al [2] found the strategy with the current smallest expected meeting time. Their strategy emerges as a result of recognizing what made Alpern's strategy successful: if the two agents do not rendezvous after their first moving pattern (1 forward, 2 backwards), then the problem simply resets because the distance between them remains to be 2 units. Their framework involves increasing the number of moves of the moving patterns and only assigning probabilities to those moving patterns such that they are all distance-preserving, which is to say that if rendezvous does not occur by the end of a cycle then the problem simply resets. In Alpern's strategy, there was only one moving pattern of length 3 that was assigned probability, but Han et al employ optimization techniques to increase the length of their moving patterns, which increases their number as well. In essence, this helps to break the symmetry of the problem. However, due to computational limitations, this was performed only up to all moving patterns of length 15, for which an expected meeting time of approximately of 4.257 was derived. Han et al use a similar technique to derive a lower bound of 4.152, and both bounds have yet to be bettered.

Both parties have also proved the following theorems: it is optimal for both agents to change direction at integer intervals and move at maximum speed throughout. Han *et al* refer to the strategies that follow these guidelines as *grid strategies*.

We would be remiss not to mention the contributions of other researchers (Anderson and Essegaier [3], Baston [4], and Uthaisombut [5]) who also found strategies that bettered Alpern's initial upper bound of 5.0.

3 The Bayesian Framework

3.1 The Essence of the Framework

Imagine that you are one of the agents attempting to rendezvous with the other agent on the line, but suddenly all information about your previous movements is destroyed. However, a genie appears to you and tells you that with probability one, the other agent is in your forward direction, and this genie orients you accordingly (that is, you know what forward they are referring to). We remark that, despite losing all of your previous information, you have everything you need to act optimally, which is to move forward with probability one ad infinitum. This seems trivial: if you know which direction the other agent is in, travelling only in that direction assures the smallest expected meeting time.

Now imagine that the same scenario with the genie occurs, except they tell you something slightly different: with probability 0.79, the other agent is in your for-

ward direction (again, you are oriented accordingly). Now, things are certainly less clear. Namely, two questions arise:

- 1. How should you act on this information so as to move optimally?
- 2. Could you perform a strategy with a smaller expected meeting time if you had more information than this?

This paper focuses on answering the first question, with the second question being the subject of future research.

3.2 Formalizing the Framework

Although the two agents' absolute starting positions will either be (-1) or (1), we are more concerned with their point of view, which is to say that each agent supposes they are starting at point (0) and that the other agent is either at (-2) or (2). Also note that neither agent is concerned with the universal notion of forward but rather their own arbitrary designation of forward. To assuage confusion, we will say that both agents think of themselves as "the agent" and they are trying to rendezvous with "the companion". We let $f(t):[0,\infty] \to \mathbb{R}$ denote the agent's position according to the agent and $g(t):[0,\infty] \to \mathbb{R}$ denote the companion's position according to the agent at time t. Notice that both f(t) and g(t) are continuous functions because no agent can jump over intervals of the line. Figure 1 is an illustrative plot of f(t) where the agent is following a grid strategy.

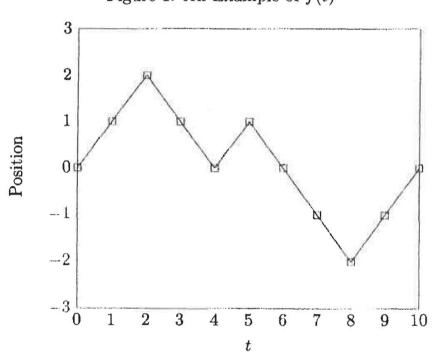


Figure 1: An Example of f(t)

Clearly, the agent knows the precise form of f(t) since it has perfect memory of what moves it has completed. If the agent knew the exact form of the function

g(t) this would be extremely helpful in expediting rendezvous since they would now know precisely in which direction to move to get closer to the companion. Although the precise form of g(t) might be unknown, there is still some useful information to be gained.

Proposition 1. The direction in which the agent is in relative to the companion does not change until rendezvous occurs.

Proof. Since both f(t) and g(t) are continuous functions, so too must the function h(t) = f(t) - g(t) be continuous. If h(t) < 0 then the agent is in the forward direction of the companion and if h(t) > 0 then the agent is in the backward direction of the companion. Suppose rendezvous has not yet occurred at time t^* , then there does not exist a $\hat{t} < t^*$ such that $h(\hat{t}) = 0$. Via the Intermediate Value Theorem, if $h(t^*) > 0$ then for all $t \le t^*$, h(t) > 0, and if $h(t^*) < 0$ then for all $t \le t^*$, h(t) < 0.

Utilizing Proposition 1, if at any time \hat{t} before rendezvous the agent is given information that $f(\hat{t}) > g(\hat{t})$, this is equivalent to being given the information that f(0) > g(0), which is to say that the agent now knows precisely the direction in which to move to rendezvous with the companion.

We take this notion and formalize it a little bit more. Let $\alpha = f(0) - g(0)$ and R be a discrete random variable that is the time until rendezvous occurs. The agent is aware that before movement begins, $P(\alpha > 0) = P(\alpha < 0) = \frac{1}{2}$. We are interested in how this probability changes according to the agent as time goes on given that the agent has so far completed some pure strategy $s \in S_t$ while following mixed strategy x (A mixed strategy is some probability measure over a set of pure strategies, which in this case are vectors of forward and backwards movements). Note that as soon as movement begins, the universe collapses on a decision: α is either greater than or less than 1, not some superposition of the two. However, the agent is aware of this reality, so instead it can compile evidence to hypothesize what it believes the value of α to be, or what the value of $P(\alpha > 0|R > t)$ is after completing s. We utilize Bayes Theorem to get an alternate expression for this probability:

$$P(\alpha > 0 | R > t) = \frac{P(R > t | \alpha > 0) P(\alpha > 0)}{P(R > t | \alpha > 0) P(\alpha > 0) + P(R > t | \alpha < 0) P(\alpha < 0)}$$

We notice that the prior $P(\alpha > 0) = P(\alpha < 0)$, which allows some factors to cancel. Additionally, although we are seeking to find $P(\alpha > 0|R > t)$, we should recognize that to do so requires three inputs: time (t), the pure strategy the agent has completed so far (s), and the mixed strategy the companion has been following so far (x). Thus, we introduce a new function $\beta(t, s, x) = P(\alpha > 0|R > t)$ since we will need to keep track of different scenarios or inputs.

$$\beta(t, s, x) = \frac{P(R > t | \alpha > 0)}{P(R > t | \alpha > 0) + P(R > t | \alpha < 0)}$$
(1)

Proposition 2. If $\beta_i(t, s, x)$ equals 1 it is optimal for the agent to move in the backward direction.

Proof. If $\beta(t, s, x) = 1$ then $P(R > t | \alpha < 0) = 0$, which means if $\alpha < 0$ then rendezvous must have already occurred, hence $\alpha > 0$ since rendezvous has not yet occurred. Thus, since $\alpha > 0$, this implies that f(0) > g(0), hence the agent should move in the backward direction to rendezvous with the companion as quickly as possible.

Clearly, calculating $\beta(t, s, x)$ has value for the agent since it could possibly reveal the optimal direction for the agent to move in. How can the agent precisely calculate $\beta(t, s, x)$? First, we note that the companion has **four** different starting configurations:

Starting Position	Forward Direction	Probability
2	1	0.25
2	-1	0.25
-2	1	0.25
-2	-1	0.25

The agent is also aware that the companion is following the same mixed strategy that it did, so the agent can assign a probability to every possible form of g(t). Thus, it will be able to calculate $P(R > t | \alpha > 0)$ and $P(R > t | \alpha < 0)$ by summing the probability that rendezvous would have occurred depending on the different forms of g(t).

Suppose that $\beta(t, s, x) \neq 1$: is there any value to this information for the agent? Hypothesizing the answer to be *yes* leads us to construct a new strategy space. We say that a *Bayesian Strategy* $(x \in B)$ is any strategy that can be constructed in the following manner:

- 1. Select a confidence function $c(\beta):[0,1]\to [0,1]$.
- 2. At time t = 0, move in the forward direction.
- 3. At time t > 0, move in the forward direction with probability $c(\beta(t, s, x))$ or in the backward direction with probability $1 c(\beta(t, s, x))$

We label each Bayesian Strategy by its confidence function, $c(\beta)$. Every Bayesian Strategy $c(\beta)$ has an equivalent mixed strategy x over all pure strategies of length n for any arbitrary n.

Suppose we select a confidence function $\hat{c}(\beta)$ and we also wish to translate it into the more familiar structure of representing some mixed strategy \hat{x} . At time t=0, the agent moves in the forward direction, to position 1, with probability one.

The agent must then decide with which probability it should move in either direction. Note that a pure strategy is represented by a vector of 1's and -1's, where

Agent's Position	Companion's Position	P(Companion's Position)
1		0.25
1	1	0.25
1	1	0.25
1	-3	0.25

a 1 is a movement in the agent's forward direction and a -1 is a movement in the agent's backward direction. First, we update \hat{x} :

\hat{x}	
Pure Strategy	Probability
[1]	1

To calculate $\beta(1,[1],\hat{x})$ all we need to know are the probabilities $P(R>t|\alpha>0)$ and $P(R>t|\alpha<0)$, which are the probabilities that rendezvous has not yet occurred given that the companion originated in the forward or backward direction respectively. If $\alpha>0$, then half of the time rendezvous will have already occurred. If $\alpha<0$, then rendezvous has not yet occurred. Thus, $P(R>t|\alpha>0)=\frac{1}{2}$ and $P(R>t|\alpha<0)=1$, so $\beta(1,[1],\hat{x})=\frac{1}{\frac{1}{2}+1}=\frac{1}{3}$. To determine the agent's next move, we input the value we just derived into $\hat{c}(\beta)$. For this example, suppose $\hat{c}\left(\frac{1}{3}\right)=0.2$. First, let's update \hat{x} :

\hat{x} Pure Strategy Probability			
[1,-1]	0.8		

Now, the situation gets more complicated because the agent has essentially branched off into two distinct realities, which means that the companion now has eight possible movements it has completed (six, if you discount the two in which it stops after one move because rendezvous has occurred). This is illustrated below:

		Companion's Position	P(Companion's Position)
2	0 2	l	0.05
2	0.2	2	0.2
2	0 2	0	0.05
2	0.2	-3	0.2
2	0 2	-4	0.05
2	0 2	-2	0.2
0	0.8	4	0.05
0	0.8	2	0.2
0	0.8	0	0.05
0	0.8	-3	0.2
0	0.8	-4	0.05
0	0.8	-2	0.2

Our next step would be to calculate $\beta(2,[1,1],\hat{x})$ and $\beta(2,[1,-1],\hat{x})$ in order to determine how the agent would move next in each of those scenarios. We continue this process ad infinitum.

Analytic Approach to Optimizing the Confidence 3.3 Function

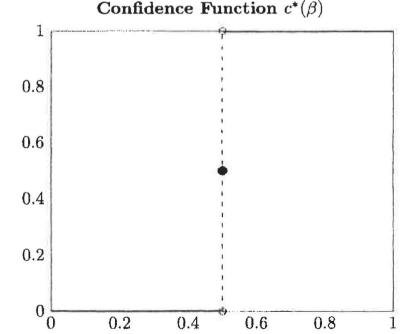
Our goal is to find the Bayesian Strategy $c(\beta)$ that yields the smallest expected meeting time. First, we establish some criteria that $c(\beta)$ must meet in order for the function to even be considered as a candidate for yielding the smallest expected meeting time:

- 1. c(1) = 1
- 2. c(0) = 0
- 3. $c(\beta): [0,1] \to [0,1]$

Next, there are features of $c(\beta)$ that we can provide intuitive justification for but have not yet proved to be necesary mathematically:

- 1. For all β , $c(\beta) = 1 c(1 \beta)$: There should be no real difference between the forward and backward direction since they are both arbitrary.
- 2. $c(\frac{1}{2}) = \frac{1}{2}$: If the agent knows it to be equally likely the companion is in either direction, moving randomly seems like a reasonable strategy
- 3. $c(\beta)$ is an non-decreasing function: If the probability the companion is in a certain direction increases, it follows that the probability the agent moves in that direction should never decrease.

There is an immediate choice for $c(\beta)$ that is easy for one to justify, which is the function below: We can think of $c^*(\beta)$ as a greedy approach: the agent is



0.4

0.6

0.8

1

0.2

always moving in the direction that it is more likely for the companion to be in. If the directions are both equally likely, the agent moves randomly.

Theorem 1. The Bayesian Strategy $c^*(\beta)$ is equivalent to Alpern's Original Strategy

Proof. First, we can express Alpern's original strategy as the following mixed strategy:

- 1. Complete the pure strategy [1, -1, -1] with probability 1.
- 2. Complete the pure strategy [1, -1, -1] with probability $\frac{1}{2}$ or [-1, 1, 1] with probability $\frac{1}{2}$.
- 3. Repeat step 2 until rendezvous

As for the Bayesian Strategy $c^*(\beta)$, the agent first moves in the forward direction, and then calculates the associated β value:

$$\beta(1, s, x) = \frac{P(R > 1 | \alpha > 0)}{P(R > 1 | \alpha > 0) + P(R > 1 | \alpha < 0)} = \frac{\frac{1}{2}}{\frac{1}{2} + 1} = \frac{1}{3}$$

Since $c^*(\frac{1}{3}) = 0$ the agent moves in the backwards direction with probability 1. Again, we calculate the β value to determine the next move:

$$\beta(2, s, x) = \frac{P(R > 2|\alpha > 0)}{P(R > 2|\alpha > 0) + P(R > 2|\alpha < 0)} = \frac{\frac{1}{2}}{\frac{1}{2} + 1} = \frac{1}{3}$$

Again, the agent moves in the backwards direction with probability 1. We notice that at this point the problem resets in that the companion is 2 units in the forward direction with probability $\frac{1}{2}$ or in the backward direction with probability $\frac{1}{2}$. Since $c(\frac{1}{2}) = \frac{1}{2}$, we can think of the agent's next movement as determining a new forward direction. Since the problem resets, we know that following $c^*(\beta)$ will now result in either completing [1, -1, -1] or [-1, 1, 1] and then resetting again. Thus, $c^*(\beta)$ reduces to the same mixed strategy as Alpern's Original Problem, so they are equivalent strategies.

Equating Alpern's Original Strategy to $c^*(\beta)$ helps provide further justification as to why it is not optimal. Clearly, the agent is behaving too bullishly when they start to compile evidence as to which direction the companion is in. Although so far we have only demonstrated that Bayesian Strategies can provide an alternative outlook on established strategies, in the next section we will show how we can derive actual bounds for any Bayesian Strategy.

4 Deriving Bounds

4.1 General Strategy

The expected meeting time given two agents are following the mixed strategy x is denoted by $\hat{T}(x,x)$. Suppose the two agents have only completed n moves, with probability p_n that rendezvous has not yet occurred. Consider the partial expected meeting time, $\hat{T}_n(x,x)$ which is the expected meeting time using the

probability of rendezvous of the first n moves and doesn't normalize the probabilities (i.e. probability p_n is "missing" from the expectation). Then the following inequality holds for all n:

$$\hat{T}(x,x) \ge \hat{T}_n(x,x) + p_n(n+1)$$

Additionally, given some mixed strategy x, we can produce an upper bound for R^s . After completing n moves, each agent retraces its steps and returns to its original position. From there, both agents restart the strategy, again retracing their steps and starting over if rendezvous does not occur after another n moves. We call the new mixed strategy produced from this procedure x^* .

$$\hat{T}(x^*, x^*) = \hat{T}_n(x, x) + p_n(\hat{T}(x^*, x^*) + 2n)$$

which implies:

$$(1-p_n)\hat{T}(x^*,x^*) = \hat{T}_n(x,x) + 2np_n$$

from which we obtain:

$$\hat{T}(x^*, x^*) = \frac{\hat{T}_n(x, x) + 2np_n}{1 - p_n} \ge R^s.$$
(2)

Although the strategy x^* ensures a finite expected meeting time, we can certainly do better. Suppose x^{**} is a strategy that is identical to x^* , except when the agents are retreating to their original position they perhaps don't retrace their steps, so there is maybe some chance the two agents rendezvous during this procedure. For example, let's say that after n moves the agent has to move forward i times and backward j times to return to where they started. With probability $\frac{1}{2}$ they move forward i times and backward j times and with probability $\frac{1}{2}$ the order is reversed. Clearly, the following inequality must hold:

$$\hat{T}(x^{**}, x^{**}) \le \hat{T}(x^{*}, x^{*}) \tag{3}$$

These bounds are useful for a few reasons. Firstly, even if $\hat{T}(x,x)$ itself is infinite, we can still derive bounds for R^s for any arbitrary n. Secondly, $\hat{T}_n(x,x)$ can be determined easily for sufficiently small n with a computer, which is useful since the closed form $\hat{T}(x,x)$ may be unknown.

4.2 Computational Approach to Optimizing the Confidence Function

We have written a program that takes as input a valid confidence function $c(\beta)$ and some number of moves n and gives as output a lower bound for $T(c(\beta), c(\beta))$ and two upper bounds for R^s , one where the agents retrace their steps as in equation 2 and one where the agents retreat in the way described before in equation 3.

To feed the program a confidence function, we first parameterize the confidence function so that its shape is completely determined when given the value of the parameter. We utilize scipy's optimize package so that we can optimize the parameter to either minimize the upper bound for R^s or maximize the lower bound

for $T(c(\beta), c(\beta))$. The goal is to find a confidence function that produces an upper bound on R^s that is better than Han *et al.*'s upper bound (approximately 4.2574). Here is the procedure we will follow:

- 1. Parameterize $c(\beta)$ so it is of some form fully determined by parameter X.
- 2. Define the function l(n) to accept as input any n and give as output the value of the minimum lower bound of $\hat{T}(c(\beta), c(\beta))$ given n moves.
- 3. Determine if there exists an n such that l(n) > 4.2574...
- 4. Define the function u(n) to accept as input any n and give as output the value of X that minimizes the best upper bound of R^s derived from $c(\beta)$ given n moves.
- 5. Select a X* that could potentially satisfy $\lim_{n\to\infty} u(n) = X^*$
- 6. Compute the best upper bound for R^s given $c(\beta)$ and X^* for as large a value of n as you would like.

Steps 2 and 3 are to check if it is possible that the specific parameterization could potentially yield a strategy that is better than the strategy of Han *et al.* Steps 5 and 6 are to actually derive a strategy and an upper bound on R^s . We will walk through this procedure completely for a parameterization of $c(\beta)$ and afterwards we will present the results in Figure 5.

Consider the following class of linear confidence functions, illustrated in Figure 2: For some $X \in [0, \infty)$,

$$c(\beta) = \begin{cases} 0 & 0 \le \beta \le \frac{1}{2} + \frac{1}{2X} \\ X(\beta - \frac{1}{2}) + \frac{1}{2} & \frac{1}{2} + \frac{1}{2X} \le \beta \le \frac{1}{2} + \frac{1}{2X} \\ 1 & \frac{1}{2} + \frac{1}{2X} \le \beta \le 1 \end{cases}$$

Figure 2: The Linear Confidence Function with parameter value $X = \sqrt{e}$

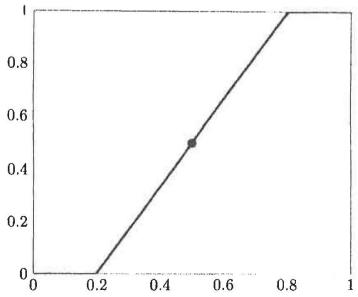


Figure 3: l(n) for the linear confidence function up to n=20.

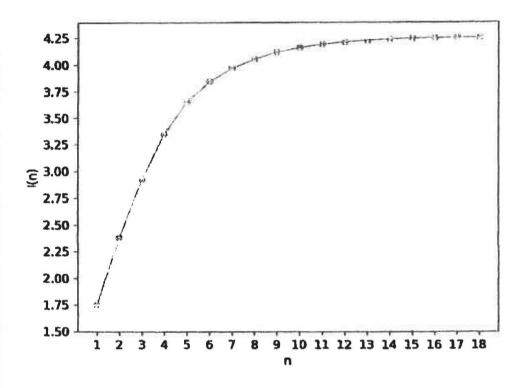
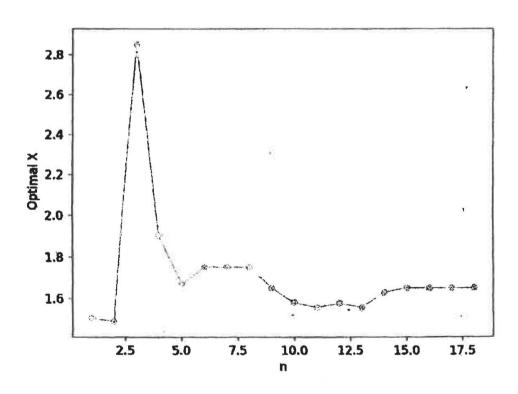


Figure 4: u(n) for the linear confidence function up to n = 18.



First, we search for an n that satisfies l(n) > 4.2574...: At n = 20, l(n) > 4.2574... so we must reject this parameterization as a candidate for besting Han et al.'s upper bound.

So now we shift gears and try to estimate $\lim_{n\to\infty} u(n)$.

It appears that \sqrt{e} is pretty close to the optimal value of X for n=18. Any selection of n will give us a valid upper bound, so now the only issue remaining is computation time. Since we have shown that this parameterization cannot best Han *et al.*'s strategy, it is not a worthy endeavor to try an extremely large n. We choose n=20 instead:

Linear Confidence Function, $X = \sqrt{e}, n = 20 \Rightarrow R^n < 4.27840$

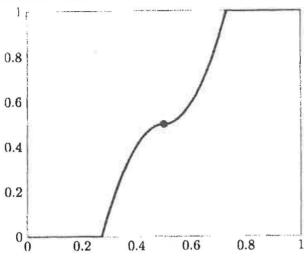
Figure 5: Results of Heuristic for Linear Confidence Function

n	Lower Bound on $T(c(\beta), c(\beta))$	Best Upper Bound on Ra
1	1.75	7.0
2	2.41275	6.58724
3	2.93635	6.29112
4	3.35853	6.12179
5	3.66280	5.73924
6	3.85250	5.15721
7	3.97808	4.93982
8	4.06390	4.75582
9	4.12475	4.66749
10	4.16590	4.56047
11	4.19474	4.49508
12	4.21480	4.42825
13	4.22907	4.39157
14	4.23893	4.35177
15	4.24573	4.32965
16	4.25039	4.31004
17	4.25363	4.29678
18	4.25586	4.28687
19	4.25739	4.28020
20	4.25847	4.27480

We also consider the following class of quadratic confidence function, illustrated in Figure 7:

For some
$$X \in [0, \infty)$$
, $c(\beta) = \begin{cases} 0 & 0 \le \beta \le \frac{1}{2} - \sqrt{\frac{1}{2X}} \\ -X(\beta - \frac{1}{2})^2 + \frac{1}{2} & \frac{1}{2} - \sqrt{\frac{1}{2X}} \le \beta \le \frac{1}{2} \\ X(\beta - \frac{1}{2})^2 + \frac{1}{2} & \frac{1}{2} \le \beta \le \frac{1}{2} + \sqrt{\frac{1}{2X}} \\ 1 & \frac{1}{2} + \sqrt{\frac{1}{2X}} \le \beta \le 1 \end{cases}$

Figure 6: The quadratic confidence function with parameter value X = 9.5



Performing the same procedure as the linear confidence function reveals that no choice of X yields an expected meeting time smaller than the strategy of Han et al. The optimal X appears to be oscillating around 9.5 at the end, which yields an upper bound of 4.32816 for n=16.

Figure 7: Results of Heuristic

Form of	$\lim_{n\to\infty} l(n)$	Estimate of		Upper
c(eta)	> 4.2574?	$\lim_{n\to\infty}u(n)$	n	Bound
Linear	Yes	√e	20	4.2748
Quadratic	Yes	9.5	16	4.32816

5 Summary

We have developed a new strategy space that takes a Bayesian approach to calculating the next move the agents should make, considering some of the information they have available to them. This provides us with an intuitive basis for deriving a strategy, as opposed to exhaustive computation. By interpreting other strategies as Bayesian ones (as we did with Alpern's original strategy) can shed light into what makes them successful or not. Additionally, we have shown how to derive both lower and upper bounds from any mixed strategy x that might only describe a finite number of moves. Finally, we have crafted a heuristic which utilizes known optimization algorithms to search for a Bayesian strategy with the smallest expected meeting time, and have derived an upper bound on R^s of 4.27840. Future work is centered around devising smarter ways to find better Bayesian strategy confidence functions.

References

- [1] S. Alpern. The rendezvous search problem. SIAM J Control Optimization, 33 (1995), 673-678.
- [2] Q. Han, D. Du, J. Vera, L. Zuluaga. Improved bounds for the symmetric rendezvous value on the line. *Oper. Res.* 56 (2008), no.3, 772-782.
- [3] Anderson, Edward J.; Essegaier, Skander. Rendezvous search on the line with indistinguishable players. SIAM J. Control Optim. 33 (1995), no. 6, 1637–1642.
- [4] Baston, Vic. Note: Two rendezvous search problems on the line. Naval Res. Logist. 46 (1999), no. 3, 335-340. 90B40 (90D05)
- [5] Uthaisombut, P. Symmetric rendezvous search on the line using moving patterns with different lengths. Working Paper, Department of Computer Science, University of Pittsburgh. (2006)