Probability in the Engineering and Informational Sciences, **30**, 2016, 576–592.
doi:10.1017/S026996481600022X

OPTIMAL STRATEGY IN "GUESS WHO?": BEYOND BINARY SEARCH

MIHAI NICA
Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012-1185, USA
E-mail: nica@cims.nyu.edu

"Guess Who?" is a popular two player game where players ask "Yes"/"No" questions to search for their opponent's secret identity from a pool of possible candidates. This is modeled as a simple stochastic game. Using this model, the optimal strategy is explicitly found. Contrary to popular belief, performing a binary search is *not* always optimal. Instead, the optimal strategy for the player who trails is to make certain bold plays in an attempt catch up. This is discovered by first analyzing a continuous version of the game where players play indefinitely and the winner is never decided after finitely many rounds.

1. INTRODUCTION

"Guess Who?" is a zero-sum two player game where both players take on a secret identity from a predefined set of characters, and the goal of the game is to be the first to discover the opponent's secret identity by process of elimination [12]. To achieve this, each player keeps track of a finite pool of possible candidates for their opponent's secret identity. Players take turns asking "Yes"/"No" question about their opponent's identity (which their opponent must answer truthfully) and use this information to reduce the pool of candidates accordingly. For example, if Player 1 asks: "Does your character have blue eyes?" and Player 2 answers "No", then Player 1 can eliminate all candidates that have blue eyes. Eventually, only one candidate remains in the pool and this last character must be their opponent's secret identity. The first player to narrow their pool down to a single candidate in this way is the winner of the game.

To model this game mathematically we will make the following assumptions:

- The secret identity of the opponent is uniformly distributed amongst all possible candidates. Because of the eliminating nature of the "Yes"/"No" questions, this property persists throughout the game. With this assumption, only the number of remaining characters in the pool is relevant to the analysis, not the details of which characters in particular are remaining.
- For any candidate pool size n and any $1 \leq b \leq n-1$, it is always possible to construct a question for which exactly b candidates correspond to the "Yes" answer. One way to do this is to sort the candidates alphabetically and ask "Does your character's name come alphabetically before __?" where the name is chosen to be the name which is b-th on the list. We use the terminology a "**bid of size** b" for such a "Yes"/"No" question.

© Cambridge University Press 2016 0269-9648/16 $25.00

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

With these assumptions, "Guess Who?" can be modeled as a so called *simple stochastic game* as defined by Condon [2]. The strategy of the game is in choosing the bid size. Players can balance risk versus reward on each turn by varying their bid size. The bid of size $b = 1$ is a risky play: it gives a small chance to win the game immediately but is more likely to reduce the candidate pool by only 1. In contrast, a bid of size $b = \lfloor \frac{1}{2} n \rfloor$ is the least risky bid. The main result of this article is to find the optimal bidding strategy for "Guess Who" and prove that it is optimal.

THEOREM 1.1: *(Optimal Strategy and Optimal Probabilities for "Guess Who?") When it is Player 1's turn, if Player 1 has n candidates in their pool and Player 2 has m candidates in their pool, then Player 1 has the following optimal strategy:*

- *If $n \geq 2^{k+1} + 1$ while $2^k + 1 \leq m \leq 2^{k+1}$ for some $k \in \mathbb{N} \cup \{0\}$, then Player 1 is in the weeds and must make a bold move to catch up! Their optimal play is a bid of size $b^*(n, m) = 2^k = 2^{\lfloor \log_2(m-1) \rfloor}$ and the probability Player 1 wins if both players play optimally is:*

$$p^\star(n, m) = \frac{2^{k+1}}{n} - \frac{2}{3} \cdot \frac{2^{2k+1} + 1}{nm}. \tag{1.1}$$

- *If $2^k + 1 \leq n \leq 2^{k+1}$ while $m \geq 2^k + 1$ for some $k \in \mathbb{N} \cup \{0\}$, then Player 1 has the upper hand and can stay ahead by making a low risk play! Their optimal play is a bid of size $b^*(n, m) = \lfloor \frac{1}{2} n \rfloor$ and the probability Player 1 wins if both players play optimally is:*

$$p^\star(n, m) = 1 - \frac{2^k}{m} + \frac{2}{3} \cdot \frac{2^{2k} + 2}{nm}. \tag{1.2}$$

The proof of Theorem 1.1 is provided in Section 4, and goes by solving a recurrence relation that $p^\star(n, m)$ satisfies. Figure 1 shows a plot of Player 1's probability of winning as a function of the pool sizes (n, m).

1.1. Bold Plays

It is an elementary entropy calculation that the low risk bid $b = \lfloor \frac{1}{2} n \rfloor$ will minimize the expected number of "Yes"/"No" questions for Player 1. This is *not* the optimal strategy in "Guess Who?" because Player 1 does not want to minimize this expected value: instead he wants to maximize the probability of reducing their pool to one person before Player 2 does. This race against the opponent is what drives the optimal bidding behavior when Player 1 is significantly behind his opponent. When this happens, Player 1's optimal strategy is a bold bid, $b = 2^{\lfloor \log_2(m-1) \rfloor}$ which depends only on Player 2's (!) remaining pool, and is always strictly $< \lfloor \frac{1}{2} n \rfloor$. This has a low probability of success (always strictly$< \frac{1}{2}$) but has the potential to put Player 1 back in the running.

The concept of risky plays with big payoffs in stochastic games has a rich history. A classic example is the single player casino game red-and-black when playing against a subfair casino studied by Dubins and Savage [3]. There is also a two player version of red-and-black, where players try to bankrupt each other, which has also been studied by Secchi [8], and adapted by Pontiggia [7]. Still more variations are studied more recently by Chen and Hsiau [1]. In all these examples, the authors find when bold plays are optimal.

Another example, more similar to "Guess Who?", is the two player dice game "Pig" where players race to 100 points while managing risk versus reward on each of their turns. Without the complication of racing against the opponent, Haigh and Roters [5] are able to conduct exact analysis. A game-show version where the players play simultaneously has

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

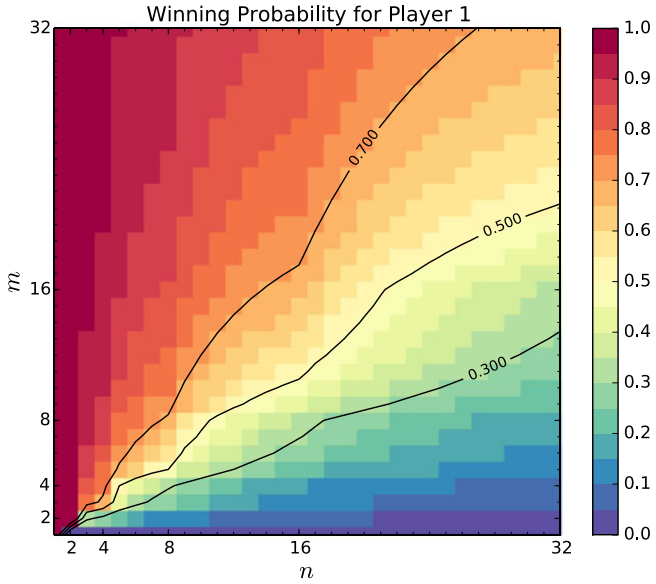


FIGURE **1.** The optimal probability $p^\star(n,m)$ for Player 1 to win the game when Player 1 has n candidates, Player 2 has m candidates, and it is Player 1's turn to guess.

also been studied by Tijms and van der Wal [10]. However, the interplay between two racing players complicates things. Neller and Presser [6] have numerically calculated the optimal strategy in situations where both players have <100 points using dynamical programming techniques, but there is no known simple description of the optimal strategy. The optimal strategy to "Guess Who?" in Theorem 1.1 is satisfying because we can find a simple exact solution on how and when to play boldly.

1.2. Log Periodicity and "Continuous Guess Who?"

The landscape of the game exhibits a log periodic behavior. In particular, for fixed $c > 0$, $p^\star(n, cn)$ does *not* converge as $n \to \infty$. Instead it approaches a function which is periodic in $\log_2(n)$. This kind of behavior is not uncommon in this kind of stochastic system; for example, see group Russian roulette studied by van de Brug, Kager, and Meester [11] or ties amongst independent and identically distributed (i.i.d.) random variables studied by Eisenberg, Stengle, and Strang [4]. In "Guess Who?" there is an asymptotic function $p^\star_\infty : \mathbb{R}^{>0} \times \mathbb{R}^{>0} \to [0,1]$ which is exactly $\log_2$-periodic in the sense that $p^\star_\infty(2x, 2y) = p^\star_\infty(x, y)$ and for which $p^\star(n,m) = p^\star_\infty(n,m) + O\left(1/nm\right)$.

This function $p^\star_\infty$ was first discovered as the probability that Player 1 wins a generalized game called "Continuous Guess Who?" which is introduced and studied in Section 3. This game is more straightforward to analyze than "Guess Who?" because there are no small number effects. Figure 2 shows a plot of $p^\star_\infty(2^k\alpha, 2^k\beta)$ when (α, β) is in the L shaped region $(2,\infty) \times (1,2] \bigcup (1,2] \times (1,\infty)$. The landscape of $p^\star_\infty$ on all of $\mathbb{R}^{>0} \times \mathbb{R}^{>0}$ can be understood by tiling scaled copies of this L shaped region onto the entire quadrant.

This log-periodicity means that if both players start with the same pool size, no matter how large, Player 1 will always have a significant first turn advantage. If he plays optimally he will win with a probability between 5/8 and 2/3, depending on the particular starting size. It would be more fair to start Player 1 with a larger pool of candidates to offset his

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

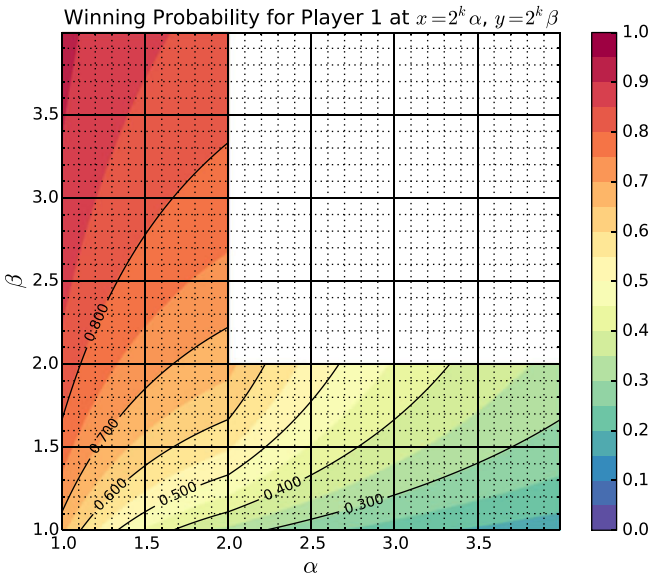


FIGURE **2.** The probability $p^\star_\infty(2^k\alpha, 2^k\beta)$ of Player 1 winning "Continuous Guess Who?" from the position $x = 2^k\alpha$, $y = 2^k\beta$ in the region $(\alpha, \beta) \in (2, \infty) \times (1, 2]$ and $(\alpha, \beta) \in (1, 2] \times (1, \infty)$. This closely approximates "Guess Who?" when n, m are large: $p^\star(n, m) = p^\star_\infty(n, m) + O(1/nm)$.

advantage from playing first. The fair way to do this is enlarge his starting pool by a factor between 4/3 and 3/2.

1.3. Real-World Applications

Even though "Guess Who?" is an abstract game, the key features of the game are more broadly applicable:

- Players are racing to reach a specific goal and get nothing for a second place finish.
- The progress of both players toward the goal is public knowledge.
- Players have the ability to balance risk versus reward on the speed of progress toward the goal.

In terms of these features, "Guess Who?" can be thought of as a toy model of some situations where two firms are competing to be the first to bring a product to market and have the ability to manage risk versus reward on the time it takes for product development. Our analysis suggests that the leading firm should take no risks (i.e. try only to minimize expected development time) and the trailing firm should take the smallest risk possible catch up to the leading firm.

An example from history is the space race between the USA and Soviet Union. After Apollo 8 became the first manned mission to orbit the Moon in December 1968, the Soviet Union went forward with a launch of its N1 rocket to attempt lunar orbit in February 1969 despite warnings from engineers on the low probability of mission success. The rocket catastrophically failed, but nevertheless the Soviet Union quickly scrambled for another attempt in July 1969, just 13 days before the launch of Apollo 11. The rocket exploded only 23 seconds after launch and resulted in the largest non-nuclear man-made explosion of all

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

time. Clearly, the Soviet Union was trying to increase the speed of development at the price of increased risk. They knew that taking risks was the best way to maximize the chance of catching up.

2. THE MODEL

DEFINITION 2.1: *Mathematician's "Guess Who?" is a game between two players played on the statespace:*

$$\mathbb{S} = \{\langle n, m, P_i\rangle \ : \ n \in \mathbb{N}, m \in \mathbb{N}, i \in \{1,2\}\}. \tag{2.1}$$

The first entry indicates the number of people remaining in Player 1's pool of possible candidates for Player 2's mystery character. (This is the number of characters who have not yet been removed from consideration through previous questions.) The second entry indicates the same thing for Player 2. The last token, either P_1 or P_2 indicates whose turn it is to play. Players alternate turns.

If the last token is P_1, it is Player 1's turn. On his turn, from the state $\langle n, m, P_1\rangle$, Player 1 makes a ***bid*** *$b_1 \in [1, n-1] \cap \mathbb{N}$. This represents a "Yes-No" question that Player 1 may ask Player 2 about his hidden character. b_1 represents the size of the set of "Yes" answers. Once Player 1 has selected his bid, the game state evolves stochastically according to the following rules:*

$$\langle n, m, P_1\rangle \text{ with Player 1 bidding } b_1 \to \begin{cases} \langle b_1, m, P_2\rangle & \text{with probability } \dfrac{b_1}{n}, \\ \langle n - b_1, m, P_2\rangle & \text{with probability } \dfrac{n-b_1}{n}. \end{cases}$$

This reflects the fact that Player 2's mystery character is uniformly distributed in the pool of characters. From this state, it is Player 2's turn. Player 2 makes a bid $b_2 \in [1, m-1] \cap \mathbb{N}$, and the state space evolves in the same way:

$$\langle n', m, P_2\rangle \text{ with Player 2 bidding } b_2 \to \begin{cases} \langle n', b, P_1\rangle & \text{with probability } \dfrac{b_2}{m}, \\ \langle n', m - b_2, P_1\rangle & \text{with probability } \dfrac{m-b_2}{m}. \end{cases}$$

All random chances in this game are independent of one another. Afterward it is Player 1's turn again and play repeats. Players continue in this way until either player reduces his pool of candidates to exactly 1. When this happens that player immediately wins[1]*. That is to say that the following states are terminal states for the game:*

$$\forall m > 1, \langle 1, m, P_2\rangle \leftrightarrow \text{Player 1 immediately wins},$$
$$\forall n > 1, \langle n, 1, P_1\rangle \leftrightarrow \text{Player 2 immediately wins}.$$

The state $\langle 1, 1, P_i\rangle$ is undefined because there is no way to reach this state without first passing through one of the previously defined terminal states.

Remark 2.2: Definition 2.1 puts the game of "Guess Who?" in the framework of a *simple stochastic game* as defined by Condon [2]. These are also sometimes called $2\frac{1}{2}$-player games,

[1] In some versions of "Guess Who?", players are required to make an additional "special" final guess, even if they have already reduced the pool to a single individual. We do not consider this rule set here.

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

where in addition to Players 1 and 2, the randomness in the game is thought of as $\frac{1}{2}$ of a player. Because of the randomness inherent in these games there is no strategy for either player that guarantees victory. (Indeed, in "Guess Who?", the opponent can always make a bid of 1, and if they are lucky, will immediately win.)

DEFINITION 2.3: *The theory of simple stochastic games (see Shapley [9] and Condon [2]) guarantee the existence of an "optimal bidding" function* $b^\star : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *and an "optimal probability" function* $p^\star : \mathbb{N} \times \mathbb{N} \to [0,1]$ *that are optimal for Player 1 in the sense that*[2]*:*

- *If Player 1 makes a bid of* $b^\star(n,m)$ *whenever the state* $\langle n, m, P_1\rangle$ *is encountered, then no matter what strategy Player 2 chooses, Player 1 wins with probability* $\geq p^\star(n,m)$*.*
- *If Player 1 uses any other bidding strategy whatsoever at the state* $\langle n, m, P_1\rangle$*, then Player 2 has a strategy that ensures that Player 1 wins with probability* $\leq p^\star(n,m)$*.*

When both players play their optimal strategies, Player 1 will win with probability exactly $p^\star(n,m)$. Since "Guess Who?" is symmetric between Player 1 and Player 2 (in the sense that the position $\langle n, m, P_1\rangle$ is identical to the position $\langle m, n, P_2\rangle$ by interchanging the names of the players), the optimal bidding function for Player 2 from $\langle n, m, P_2\rangle$ is $b^\star(m,n)$ and Player 2's optimal probability to win from $\langle n, m, P_2\rangle$ is $p^\star(m,n)$. By the same token, Player 1's probability to win from $\langle n, m, P_2\rangle$ is $1 - p^\star(m,n)$ since Player 1 wins if and only if Player 2 loses. Thus, both players strategies and bids can be encapsulated in a single function. Without loss of generality then, we will thus normally take the point of view of Player 1 in our analysis. Because of this symmetry, the optimal probability function satisfies a nice recurrence relation.

PROPOSITION 2.4: *The functions* $p^\star(n,m)$ *and* $b^\star(n,m)$ *satisfy the following recurrence relation:*

$$p^\star(n,m) = \max_{b\in[1,n-1]} \left\{ 1 - \frac{b}{n} p^\star(m,b) - \frac{n-b}{n} p^\star(m,n-b) \right\}, \tag{2.2}$$

$$b^\star(n,m) = \arg\max_{b\in[1,n-1]} \left\{ 1 - \frac{b}{n} p^\star(m,b) - \frac{n-b}{n} p^\star(m,n-b) \right\}. \tag{2.3}$$

(For the case the argmax is not unique: any b *that maximizes the function will work for Player 1's optimal strategy.)*

PROOF: Let $p_b(n,m)$ be the probability that Player 1 wins from the position $\langle n, m, P_1\rangle$ if he bids b at $\langle n, m, P_1\rangle$ and both players play optimally thereafter. By the rules of the game $p_b(n,m) = (b/n)\,(1 - p^\star(m,n)) + (n - b/n)\,(1 - p^\star(m,n-b))$ since with probability b/n we move to the position $\langle b, m, P_2\rangle$ where Player 1's probability to win is $1 - p^\star(m,n)$ and with probability $n - b/n$ we move to the position $\langle n-b, m, P_2\rangle$ where Player 1's probability to win is $1 - p^\star(m,n-b)$.

It must be that $p^\star(n,m) = \max_{b\in[1,n-1]} p_b(n,m)$ since Player 1 must bid some value of b at the position $\langle n, m, P_1\rangle$, and the optimal strategy is no worse than any fixed bid b. Similarly, any $b^\star$ that has $p_{b^\star}(n,m) = p^\star(n,m)$ means that bidding $b^\star$ at the position $\langle n, m, P_1\rangle$ is an optimal bid. ■

[2] This function $p^\star(n,m)$ is sometimes called the *value* of the node $\langle n, m, P_1\rangle$ and the strategy $b^\star$ is called a *positional* strategy since it depends only on the current position.

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

Remark 2.5: "Guess Who?" has a nice structure because the sum $s = n + m$ is a *strictly decreasing* function as the game progresses. Indeed, this sum is guaranteed to decrease by at least 1 on each players turn. This observation, along with the initial data $\forall m > 1$, $p^\star(1, m) = 1$ and $\forall n > 1$, $p^\star(n, 1) = 0$, mean that $p^\star(n, m)$, $b^\star(n, m)$ can be explicitly computed inductively by the very simple Algorithm 1 given below[3]. At each stage s, the algorithm has already computed $p^\star(n, m)$ for all pairs (n, m) with $n + m \leq s - 1$. The values $p^\star(m, b)$ and $p^\star(m, n - b)$, which appear in the maximization have $m + b \leq s - 1$ and $m + (n - b) \leq s - 1$ since $1 \leq b \leq n - 1$ and $m + n = s$, and are hence already known. The proof of the main Theorem 1.1 also uses this observation as part of a proof by induction. The output when $r = 64$ is displayed in Figure 1 for $1 \leq n, m \leq 32$.

Algorithm 1 Numerically computing $p^\star(n, m)$ and $b^\star(n, m)$.

Input: $r \geq 3$ *Output:* The value $p^\star(n, m)$ and $b^\star(n, m)$ for all pairs (n, m) with $3 \leq n + m \leq r$

0. Initialize $p^\star(1, m) = 1$ for all $m > 1$ and $p^\star(n, 1) = 0$ for all $n > 1$
1. **for** $s = 3$ to r:
2. **for** (n, m) with $n + m = s$:
3. $p^\star(n, m) \leftarrow \max_{b \in [1, n-1]} \left\{1 - \frac{b}{n} p^\star(m, b) - \frac{n-b}{n} p^\star(m, n - b)\right\}$
4. $b^\star(n, m) \leftarrow \operatorname{argmax}_{b \in [1, n-1]} \left\{1 - \frac{b}{n} p^\star(m, b) - \frac{n-b}{n} p^\star(m, n - b)\right\}$

2.1. "In the weeds" and "The upper hand"

We now divide the statespace $\mathbb{S}$ into some subsets which turn out to be relevant for discussing the optimal strategy in the game. Except for the trivial states where $n = 1$ or $m = 1$, these subsets partition the entire statespace $\mathbb{S}$ in the sense that each element of the statespace is in exactly one of the subsets. This division was discovered by careful examination of the output of Algorithm 1.

Definition 2.6: *For $k \in \mathbb{N} \cup \{0\}$, define the set $W_{k,P_1} \subset \mathbb{S}$ by:*

$$W_{k,P_1} := \left\{ \langle n, m, P_1 \rangle \ : \ 2^{k+1} < n \text{ and } 2^k < m \leq 2^{k+1} \right\}. \tag{2.4}$$

*When $\langle n, m, P_1 \rangle \in W_{k,P_1}$, we say that **Player 1 is in the weeds at level** k, or if $\langle n, m, P_1 \rangle \in \bigcup_{k=0}^{\infty} W_{k,P_1}$ we say **Player 1 is in the weeds** without specifying which level. Similarly, $W_{k,P_2} := \left\{ \langle n, m, P_2 \rangle \ : \ 2^{k+1} < m \text{ and } 2^k < n \leq 2^{k+1} \right\}$ and we use the same terminology for Player 2.*

Definition 2.7: *For $k \in \mathbb{N} \cup \{0\}$, define the set $U_{k,P_1} \subset \mathbb{S}$ by:*

$$U_{k,P_1} := \left\{ \langle n, m, P_1 \rangle \ : \ 2^k < n \leq 2^{k+1} \text{ and } 2^k < m \right\}. \tag{2.5}$$

*When $\langle n, m, P_1 \rangle \in U_{k,P_1}$, we say that **Player 1 has the upper hand at level** k, or if $\langle n, m, P_1 \rangle \in \bigcup_{k=0}^{\infty} U_{k,P_1}$ we say **Player 1 has the upper hand** without specifying which level. Similarly, $U_{k,P_2} := \left\{ \langle n, m, P_2 \rangle \ : \ 2^k < m \leq 2^{k+1} \text{ and } 2^k < n \right\}$ and we use the same terminology for Player 2.*

[3] Note that for some (n, m) the value $b^\star(n, m)$ is not unique and will depend on how argmax is implemented.

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

With these two definitions solidified, the description of the optimal strategy for the game is very simple: players should bid 2^k when in the weeds at level k and should bid $\lfloor \frac{1}{2}n \rfloor$ when they have the upper hand.

3. "CONTINUOUS GUESS WHO?"

The exact formula in Theorem 1.1 is proven by induction, which does does not shed much light on how the formula is discovered or what the individual terms in it represent. To aid understanding, we present the original method that led to the formula $p^\star(n,m)$. From numerically computing $p^\star(n,m)$ and $b^\star(n,m)$ for some small values of n,m using Algorithm 1, we get an ansatz on the optimal bidding strategy $b^\star(n,m)$. We then exactly calculate the probability for either player to win when they follow this bidding ansatz in a modified version of "Guess Who?" which is defined in such a way to remove small number effects. The probability of winning in this modified game turns out to be a good approximation to the probability of winning in "Guess Who?" in the asymptotic regime when the pool sizes n,m are very large and a small correction to this formula gives the exact formula $p^\star(n,m)$.

DEFINITION 3.1: *"Continuous Guess Who?" is a two player game on the statespace:*

$$\mathbb{S}_\infty = \left\{ \langle x, y, P_i \rangle \ : \ x \in \mathbb{R}^{>0}, y \in \mathbb{R}^{>0}, i \in \{1,2\} \right\}. \tag{3.1}$$

The partition of the statespace into "in the weeds" and "the upper hand" are generalized to this continuous setting by allowing any exponent $k \in \mathbb{Z}$:

$$W^\infty_{k,P_1} := \left\{ \langle 2^k\alpha, 2^k\beta, P_1 \rangle \ : \ \alpha \in (2,\infty) \text{ and } \beta \in (1,2] \right\} \subset \mathbb{S}_\infty, k \in \mathbb{Z}, \tag{3.2}$$

$$U^\infty_{k,P_1} := \left\{ \langle 2^k\alpha, 2^k\beta, P_1 \rangle \ : \ \alpha \in (1,2] \text{ and } \beta \in (1,\infty) \right\} \subset \mathbb{S}_\infty, k \in \mathbb{Z}. \tag{3.3}$$

W^∞_{k,P_2}, U^∞_{k,P_2}, $k \in \mathbb{Z}$ are defined analogously. Play proceeds in a similar way to "Guess Who?": on Player 1's turn, from the state $\langle x, y, P_1 \rangle$, Player 1 is allowed to make any (not necessarily an integer) size bid $b_1 \in (0,x)$ and the game state evolves stochastically according to:

$$\text{Starting at } \langle x, y, P_1 \rangle \text{ with Player 1 bidding } b_1 \to \begin{cases} \langle b_1, y, P_2 \rangle & \text{with probability } \dfrac{b_1}{x}, \\ \langle x - b_1, y, P_2 \rangle & \text{with probability } \dfrac{x - b_1}{x}. \end{cases}$$

On Player 2's turn he makes a bid $b_2 \in (0,y)$ and the game state evolves analogously. All of these random outcomes are independent.

The game has no ending point: players continue playing indefinitely. Instead, the events $\{$Player 1 loses$\}$, $\{$Player 2 loses$\}$, $\{$The game is a draw$\}$ are defined abstractly as events on the underlying probability space. The event that Player 1 loses is defined to be the event that Player 1 is in the weeds "almost always"; that is there exists some time beyond which Player 1 is trapped in the weeds forevermore. Player 2 losing is defined analogously. The event that the game is a draw is defined to be the event that both players each alternate being in the weeds infinitely often. Denoting the state of the game after t steps as $S_t \in \mathbb{S}_\infty$,

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

this is precisely:

$$\{\textit{Player 1 loses}\} := \left\{ \exists T \in \mathbb{N} \textit{ such that } \forall t > T, S_t \in \bigcup_{k=-\infty}^{\infty} W_{k,1}^{\infty} \textit{ on Player 1's turns} \right\}$$

$$\{\textit{Player 2 loses}\} := \left\{ \exists T \in \mathbb{N} \textit{ such that } \forall t > T, S_t \in \bigcup_{k=-\infty}^{\infty} W_{k,2}^{\infty} \textit{ on Player 2's turns} \right\}$$

$$\{\textit{The game is a draw}\} := \left\{ S_t \in \bigcup_{k=-\infty}^{\infty} W_{k,1}^{\infty} \textit{ i.o. and } S_t \in \bigcup_{k=-\infty}^{\infty} W_{k,2}^{\infty} \textit{ i.o.} \right\}.$$

We also set $\{\textit{Player 2 wins}\} := \{\textit{Player 1 loses}\}$ *and* $\{\textit{Player 1 wins}\} := \{\textit{Player 2 loses}\}$.

Remark 3.2: "Continuous Guess Who?" is a very peculiar game because of the strange way that winning is defined. Indeed, the players cannot know if they have won or lost the game after finitely many rounds of the game. (There is always a positive probability the trailing player will recover and win, no matter how dire their situation may become.) Practically speaking, one should think of "Continuous Guess Who?" only as a tool that closely models ordinary "Guess Who?" when the length of the game is very long. In the remaining analysis in this section, we will show that when both players play according to an optimal bidding ansatz, the probabilities of winning from any state can be calculated exactly and the probability of a draw is 0.

Definition 3.3 Optimal Bidding Ansatz: *We assume both players play according to the following strategy:*

- *When Player 1 is in the weeds at level* k, $\left\langle 2^k\alpha, 2^k\beta, P_1 \right\rangle \in W_{k,P_1}^{\infty}$, *Player 1 makes a bid of* 2^k.
- *When Player 1 has the upper hand at level* k, $\langle x, y, P_1 \rangle \in U_{k,P_1}^{\infty}$, *Player 1 makes a bid of* $\frac{1}{2}x$.

Player 2 plays analogously. This ansatz was hypothesized by examining output from Algorithm 1.

Definition 3.4: *Define* $p_{\infty}^{\star} : \mathbb{R}^{>0} \times \mathbb{R}^{>0} \to (0,1)$ *to be the probability that Player 1 wins starting from the position* $\langle x, y, P_1 \rangle$ *on his turn when both players play according to the optimal bidding ansatz:*

$$p_{\infty}^{\star}(x,y) := \mathbf{P}\left(\textit{Player 1 wins from } \langle x, y, P_1 \rangle\right). \tag{3.4}$$

Lemma 3.5: *For any* $j \in \mathbb{Z}$, $p_{\infty}^{\star}(x,y)$ *satisfies:*

$$p_{\infty}^{\star}(x,y) = p_{\infty}^{\star}\left(2^j x, 2^j y\right). \tag{3.5}$$

Proof: Define the map $\phi : \langle x, y, P_i \rangle \to \left\langle 2^j x, 2^j y, P_i \right\rangle$ and we will show that the winning probability $p_{\infty}^{\star}$ is invariant under the map ϕ. Notice firstly that ϕ sends the sets $W_{k,P_i}^{\infty} \to W_{k+j,P_i}^{\infty}$ and $U_{k,P_i}^{\infty} \to U_{k+j,P_i}^{\infty}$. Thus, this map has no effect on whether Player 1 is in the weeds or has the upper hand. We next observe that under the optimal bidding ansatz, the probability of a bid being a success is also invariant under the map ϕ. Indeed, when Player 1

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

has the upper hand, he makes a no-risk bid and his pool size is plainly halved; when Player 1 is in the weeds at level k, $\langle 2^k\alpha, 2^k\beta, P_1\rangle \in W^{\infty}_{k,P_1}$, $\alpha > 2$, $\beta \in (1,2]$, he makes a bid of size 2^k and his pool size after the bid is:

$$\begin{cases} 2^k & \text{with probability } \dfrac{2^k}{2^k\alpha}, \\ 2^k\alpha - 2^k & \text{with probability } \dfrac{2^k\alpha - 2^k}{2^k\alpha}. \end{cases}$$

The factor 2^k cancels, and this probability depends only on α, which is invariant under ϕ. From the above considerations we see that if we let S_t be the state after t rounds started from $\langle x, y, P_1\rangle$ and we let R_t to be the state after t rounds started from $\langle 2^j x, 2^j y, P_1\rangle$, then we have equality in distribution:

$$R_t \overset{d}{=} \phi(S_t).$$

From this invariance, along with the fact that ϕ does not change whether or not Player 1 is in the weeds, and the very definition of Player 1 winning the game, we conclude that $p^{\star}_{\infty}$ is invariant under the map ϕ as desired. ■

LEMMA 3.6: *Assume both players play "Continuous Guess Who?" according to the optimal bidding ansatz. Suppose we start from the position $\langle 2^k\alpha, 2^k\beta, P_1\rangle$ with $\alpha > 2$ and $\beta \in (1,2]$, so that Player 1 is in the weeds at level $k \in \mathbb{Z}$. Then the probability that Player 1 ever exits the weeds at any point throughout the game is $2/\alpha$.*

PROOF: Since Player 1 is in the weeds at level k, his first bid is 2^k. This means that his pool size becomes:

$$\begin{cases} 2^k & \text{with probability } \dfrac{2^k}{2^k\alpha}, \\ 2^k\alpha - 2^k & \text{with probability } \dfrac{2^k\alpha - 2^k}{2^k\alpha}. \end{cases}$$

If his pool size becomes 2^k, the pool sizes are now $x = 2^k$, $y \geq 2^{k+1}$ meaning that Player 1 now has escaped the weeds and has the upper hand. Thus, Player 1 escapes from the weeds on the first guess with probability $2^k/2^k\alpha = 1/\alpha$.

If Player 1 fails to escape the weeds on this turn, his new pool size is $2^k\alpha - 2^k = 2^{k-1}(2(\alpha-1))$. At this stage Player 2 has the upper hand, and will make a bid of exactly $(1/2)2^k\beta$ and so Player 2's new pool size is halved and will be exactly $2^{k-1}\beta$. Thus, if Player 1 fails to escape the weeds at this level, the game state moves to $x = 2^{k-1}(2(\alpha-1))$ and $y = 2^{k-1}\beta$. By Lemma 3.5 the probability of winning is invariant under changing k, and so this position is handled by the same analysis as above with a new α-value which is $\alpha' = 2(\alpha-1)$. Player 1's probability to escape the weeds at this stage is $\frac{1}{\alpha'} = \frac{1}{2(\alpha-1)}$. This analysis can be repeated over and over again. Finally, to find the probability Player 1 ever escapes the weeds, we sum up his probability to escape at each stage:

$$\begin{aligned} &\mathbf{P}(\text{Player 1 ever escapes the weeds}) \\ &= \frac{1}{\alpha} + \left(1 - \frac{1}{\alpha}\right)\left(\frac{1}{2(\alpha-1)} + \left(1 - \frac{1}{2(\alpha-1)}\right)\left(\frac{1}{2\left(2(\alpha-1)-1\right)} + \cdots\right)\right). \end{aligned}$$

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

Fortunately this expansion telescopes and we remain with:

$$\begin{aligned}&\mathbf{P}\,(\text{Player 1 ever escapes the weeds})\\&= \frac{1}{\alpha} + \left(\frac{\alpha-1}{\alpha}\right)\frac{1}{2(\alpha-1)} + \left(\frac{\alpha-1}{\alpha}\right)\left(\frac{2(\alpha-1)-1}{2(\alpha-1)}\right)\left(\frac{1}{2(2(\alpha-1)-1)}\right) + \cdots\\&= \frac{1}{\alpha} + \frac{1}{2\alpha} + \frac{1}{4\alpha} + \cdots.\end{aligned}$$

Evaluating this infinite sum gives the desired result of $2/\alpha$. ■

Next we prove a lemma about the position $\langle 2^k \cdot 4, 2^k \cdot 2, P_1\rangle$. This position is important because it is naturally encountered when both players play according to the optimal bidding ansatz.

LEMMA 3.7: *Assume both players play "Continuous Guess Who?" according to the optimal bidding ansatz. Starting from the position* $\langle 2^k \cdot 4, 2^k \cdot 2, P_1\rangle$ *(i.e.* $\alpha = 4, \beta = 2$*), the probability Player 1 loses the game is* $2/3$ *and the probability of a draw is* 0.

PROOF: Player 1 can lose the game in one of two disjoint ways:

(a) Player 1 never gets out of the weeds. The probability that Player 1 never gets out of the weeds is $1 - (2/4)$ by Lemma 3.6.

(b) Player 1 successfully gets out of the weeds at some point and puts Player 2 in the weeds; then Player 2 himself gets out of the weeds; and finally Player 1 subsequently loses the game from that position. The probability of this chain of events is calculated as follows: by Lemma 3.6, with probability $\frac{2}{4}$, Player 1 does get out of the weeds at some point, and the position will then be $x = 2^\ell \cdot 2$ and $y = 2^\ell \cdot 4$ for some ℓ, (i.e. the game evolves to $\alpha' = 2, \beta' = 4$). The position must be of this form because Player 2's total is always a power of 2 up to this point (he started at a power of 2 and is always halving his pool up to this point) and Player 1's bid is the largest power of 2 which is strictly less than Player 2's pool size. From this new position, the probability that Player 2 ever gets out of the weeds from this point is $\frac{2}{4}$ by Lemma 3.6. If Player 2 does successfully get out of the weeds, the position will be $x = 2^j \cdot 4$ and $y = 2^j \cdot 2$ for some j by the same reasoning as above (i.e. we the state has evolved back to the original position $\alpha'' = 4, \beta'' = 2$). Thus, we have:

$$\begin{aligned}&\mathbf{P}\left(\text{P1 loses from } \left\langle 2^k \cdot 4, 2^k \cdot 2, P_1\right\rangle\right)\\&\quad= \mathbf{P}\,(\text{P1 never escapes}) + \mathbf{P}\,(\text{P1 escapes})\,\mathbf{P}\,(\text{P2 escapes})\,\mathbf{P}\left(\text{P1 loses from } \langle 2^{j}\cdot 4, 2^{j}\cdot 2, P_1\rangle\right)\\&\quad= \left(1 - \frac{2}{4}\right) + \frac{2}{4}\cdot\frac{2}{4}\cdot\mathbf{P}\left(\text{P1 loses from } \left\langle 2^j \cdot 4, 2^j \cdot 2, P_1\right\rangle\right)\\&\quad= \frac{1}{2} + \frac{1}{4}\mathbf{P}\left(\text{P1 loses from } \left\langle 2^j \cdot 4, 2^j \cdot 2, P_1\right\rangle\right).\end{aligned}$$

By Lemma 3.5, the positions $\langle 2^k \cdot 4, 2^k \cdot 2, P_1\rangle$ and $\langle 2^j \cdot 4, 2^j \cdot 2, P_1\rangle$ are identical with regard to losing probability. Solving the linear equation $p = (1/2) + (1/4)p$ gives $p = 2/3$, as desired. The probability of a draw is 0 because the probability of switching from one player being in the weeds to the other being in the weeds from

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

this position has probability 2/4 each time, and so having infinity many switches is a probability 0 event by the Borel–Cantelli lemma. ■

PROPOSITION 3.8: *Assume both players play "Continuous Guess Who?" according to the optimal bidding ansatz. Suppose we are at the position* $\langle x, y, P_1\rangle = \langle 2^k\alpha, 2^k\beta, P_1\rangle$ *with* $\alpha > 2$ *and* $\beta \in (1, 2]$*, so that Player 1 is in the weeds at level* $k \in \mathbb{Z}$*. Then the probability of a draw is* 0 *and the probability that Player 1 wins the game is:*

$$p^\star_\infty(x, y) = \frac{2^{k+1}}{x} - \frac{2}{3}\frac{2^{2k+1}}{xy}. \tag{3.6}$$

PROOF: We first calculate the probability that Player 1 loses. As in Lemma 3.7, Player 1 can lose the game in one of two disjoint ways:

(a) Player 1 never gets out of the weeds. The probability that Player 1 never gets out of the weeds is $1 - (2/\alpha)$ by Lemma 3.6.

(b) Player 1 successfully gets out of the weeds at some point and puts Player 2 in the weeds; then Player 2 himself gets out of the weeds; and finally player Player 1 subsequently loses the game from that position. If Player 1 does get out of the weeds, the position will be $x = 2^\ell \cdot 2$ and $y = 2^\ell \cdot (2\beta)$ for some ℓ, (i.e. the position evolves to $\alpha' = 2, \beta' = 2\beta$). Thus, the probability that Player 2 gets out of the weeds from this point is $2/2\beta$ by Lemma 3.6. If Player 2 does successfully get out of the weeds, the position will be $x = 2^j \cdot 4$ and $y = 2^j \cdot 2$ for some j (i.e. we will be at $\alpha'' = 4, \beta'' = 2$). We have already analyzed this position in Lemma 3.7. Thus, we have:

$$\begin{aligned}
&\mathbf{P}\left(\text{P1 loses from } \left\langle 2^k\alpha, 2^k\beta, P_1\right\rangle\right)\\
&\quad= \mathbf{P}\left(\text{P1 never escapes}\right) + \mathbf{P}\left(\text{P1 escapes}\right)\mathbf{P}\left(\text{P2 escapes}\right)\mathbf{P}\left(\text{P1 loses from } \left\langle 2^{j}\cdot 4, 2^{j}\cdot 2, P_1\right\rangle\right)\\
&\quad= \left(1 - \frac{2}{\alpha}\right) + \frac{2}{\alpha}\left(\frac{2}{2\beta}\right)\mathbf{P}\left(\text{P1 loses from } \left\langle 2^j\cdot 4, 2^j\cdot 2, P_1\right\rangle\right)\\
&\quad= \left(1 - \frac{2}{\alpha}\right) + \frac{2}{\alpha}\left(\frac{2}{2\beta}\right)\frac{2}{3}.
\end{aligned}$$

The probability of a draw is 0 since if more than two changes of the upper hand occur, the game passes through the position $\langle 2^j \cdot 4, 2^j \cdot 2, P_1\rangle$, from which the probability of a draw is 0 by Lemma 3.7. Finally then, taking the complement and plugging in $x = 2^k\alpha, y = 2^k\beta$ gives the result. ■

PROPOSITION 3.9: *Assume both players play "Continuous Guess Who?" according to the optimal bidding ansatz. Suppose we are at the position* $\langle x, y, P_1\rangle = \langle 2^k\alpha, 2^k\beta, P_1\rangle$ *with* $\alpha \in (1, 2]$ *and* $\beta > 2$*, so that Player 1 has the upper hand at level* $k \in \mathbb{Z}$*. Then, the probability of a draw is 0 and the probability that Player 1 wins the game is:*

$$p^\star_\infty(x, y) = 1 - \frac{2^k}{y} + \frac{2}{3}\frac{2^{2k}}{xy}. \tag{3.7}$$

PROOF: When Player 1 has the upper hand, he bids exactly $(1/2)2^k\alpha$ and puts Player 2 in the weeds. By the symmetry between Players 1 and 2, we use Proposition 3.8 to see the

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

probability of a draw is 0 and to compute:

$$\begin{aligned}
&\mathbf{P}\left(\text{Player 1 wins from state}\left\langle 2^k\alpha, 2^k\beta, P_1\right\rangle\right)\\
&\quad=\mathbf{P}\left(\text{Player 1 wins from state}\ \left\langle 2^{k-1}\alpha, 2^{k-1}\left(2\beta\right), P_2\right\rangle\right)\\
&\quad=1-\mathbf{P}\left(\text{Player 2 wins from state}\ \left\langle 2^{k-1}\alpha, 2^{k-1}\left(2\beta\right), P_2\right\rangle\right)\\
&\quad=1-\mathbf{P}\left(\text{Player 1 wins from state}\ \left\langle 2^{k-1}\left(2\beta\right), 2^{k-1}\alpha, P_1\right\rangle\right)\\
&\quad=1-\left(\frac{2}{2\beta}-\frac{2}{2\beta}\left(\frac{2}{2\alpha}\right)\frac{2}{3}\right).
\end{aligned}$$

Plugging in $x = 2^k\alpha$,$y = 2^k\beta$ gives the result. ■

Remark 3.10: Despite the difference between "Continuous Guess Who?" and ordinary "Guess Who?", this analysis comes remarkably close to the true value $p^\star(n,m)$ when evaluated at integer values. Indeed we have:

$$p^\star(n,m)-p^\star_\infty(n,m)=\begin{cases}-\dfrac{2}{3}\dfrac{1}{nm} & \text{if } \langle n,m,P_1\rangle\in W_{k,P_1},\\[2ex] \dfrac{4}{3}\dfrac{1}{nm} & \text{if } \langle n,m,P_1\rangle\in U_{k,P_1}.\end{cases} \tag{3.8}$$

The corrections $-(2/3)(1/nm)$ and $(4/3)(1/nm)$ were found serendipitously by comparing numerical solutions of $p^\star$ to the exact formula $p^\star_\infty$. The exact formula for $p^\star$ was first hypothesized by adding these simple errors onto $p^\star_\infty$. A plot of $p^\star_\infty$ is provided in Figure 2.

4. PROOF OF THEOREM 1.1

Having guessed at the formula using the above analysis, we now prove Theorem 1.1 rigorously. Define functions:

$$q_{W_k}(n,m):=\frac{2^{k+1}}{n}-\frac{2}{3}\frac{2^{2k+1}+1}{mn}, \tag{4.1}$$

$$q_{U_k}(n,m):=1-\frac{2^k}{m}+\frac{2}{3}\frac{2^{2k}+2}{mn}. \tag{4.2}$$

By stitching these functions together, these define the total function $q(n,m)$:

$$q(n,m)=\begin{cases}q_{W_k}(n,m) & \text{if } \langle n,m,P_1\rangle\in W_{k,P_1},\\ q_{U_k}(n,m) & \text{if } \langle n,m,P_1\rangle\in U_{k,P_1},\\ 1 & \text{if } n=1,\\ 0 & \text{if } m=1.\end{cases} \tag{4.3}$$

LEMMA 4.1: *Suppose that* $\langle n,m,P_1\rangle\in W_{k,P_1}$ *for some* k. *Then:*

$$q_{W_k}(n,m)=\max_{b\in[1,n-1]}\left\{1-\frac{b}{n}q(m,b)-\frac{n-b}{n}q(m,n-b)\right\}, \tag{4.4}$$

and the maximum is achieved at $b=2^k$.

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

PROOF: By symmetry between b and $n-b$, we assume without loss that $1 \le b \le \lfloor \frac{1}{2}n \rfloor$. We will show that $1-(b/n)q(m,b)-(n-b/n)q(m,n-b) \le q_{W_k}(n,m)$ for all possible bids b with equality achieved at $b=2^k$. Consider the following cases:

Case I: $\langle b,m,P_2\rangle \in W_{\ell,P_2}$ with $2^\ell < b \le 2^{\ell+1}$ and $\langle n-b,m,P_2\rangle \in U_{k,P_2}$.

We begin by simplifying:

$$\begin{aligned} &1-\frac{b}{n}q_{W_\ell}(m,b)-\frac{n-b}{n}q_{U_k}(m,n-b) \\ &\quad = 1-\frac{b}{n}\left(\frac{2^{\ell+1}}{m}-\frac{2}{3}\frac{2^{2\ell+1}+1}{mb}\right)-\frac{n-b}{n}\left(1-\frac{2^k}{n-b}+\frac{2}{3}\frac{2^{2k}+2}{m(n-b)}\right) \\ &\quad = \frac{b}{n}\left(1-\frac{2^{\ell+1}}{m}\right)+\frac{2^k}{n}+\frac{2}{3}\frac{2^{2\ell+1}-2^{2k}-1}{mn}. \end{aligned} \tag{4.5}$$

Since $\langle b,m,P_2\rangle \in W_{\ell,P_2}$, we have $2^{\ell+1} < m$ here. It follows that for a fixed ℓ, the above is a monotone increasing function of b. We thus obtain an inequality by replacing b by its maximum value $b=2^{\ell+1}$, with equality if and only if $b=2^{\ell+1}$:

$$\begin{aligned} &1-\frac{b}{n}q_{W_\ell}(m,b)-\frac{n-b}{n}q_{U_k}(m,n-b) \\ &\quad \le \frac{2^{\ell+1}}{n}\left(1-\frac{2^{\ell+1}}{m}\right)+\frac{2^k}{n}+\frac{2}{3}\frac{2^{2\ell+1}-2^{2k}-1}{mn} \\ &\quad = \frac{2^{\ell+1}}{n}\left(1-\frac{2}{3}\frac{2^{\ell+1}}{m}\right)+\frac{2^k}{n}-\frac{2}{3}\frac{2^{2k}+1}{mn}. \end{aligned}$$

This is monotone increasing as a function of $2^{\ell+1}$ since $2^{\ell+1} < m$. To see this, notice that if we increment ℓ by one, the difference is:

$$2^{\ell+1}\left(1-\frac{2}{3m}2^{\ell+1}\right)-2^\ell\left(1-\frac{2}{3m}2^\ell\right)=2^\ell\left(1-\frac{2^{\ell+1}}{m}\right)>0.$$

Thus, this function is maximized when ℓ takes its maximum value. We notice that we must have $\ell \le k-1$, since $\langle b,m,P_2\rangle \in W_{\ell,P_2}$ gives $2^{\ell+1} < m$ and $\langle n,m,P_1\rangle \in W_{k,P_1}$ gives $m \le 2^{k+1}$. Thus, this function is maximized when $\ell = k-1$, with equality if and only if $\ell = k-1$. By inspection, plugging in $\ell = k-1, b=2^k$ gives the maximum value is exactly $q_{W_k}(n,m)$ as desired.

Case II: $\langle b,m,P_2\rangle \in U_{k,P_2}$ and $\langle n-b,m,P_2\rangle \in U_{k,P_2}$.

Consider here:

$$\begin{aligned} &1-\frac{b}{n}q_{U_k}(m,b)-\frac{n-b}{n}q_{U_k}(m,n-b) \\ &\quad = 1-\frac{b}{n}\left(1-\frac{2^k}{b}+\frac{2}{3}\frac{2^{2k}+2}{mb}\right)-\frac{n-b}{n}\left(1-\frac{2^k}{n-b}+\frac{2}{3}\frac{2^{2k}+2}{m(n-b)}\right) \\ &\quad = 0+\frac{2^k}{n}-\frac{2}{3}\frac{2^{2k}+2}{mn}+\frac{2^k}{n}-\frac{2}{3}\frac{2^{2k}+2}{mn} \\ &\quad = \frac{2^{k+1}}{n}-\frac{2}{3}\frac{2^{2k+1}+4}{mn} \\ &\quad < \frac{2^{k+1}}{n}-\frac{2}{3}\frac{2^{2k+1}+1}{mn} = q_{W_k}(n,m). \end{aligned}$$

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

Interestingly, this does not depend on b at all, and is always worse than $q_{W_k}(n,m)$ by $2/mn$.

Case III: $\langle b, m, P_2\rangle \in W_{\ell_b,P_2}$ and $\langle n-b, m, P_2\rangle \in W_{\ell_{n-b},P_2}$, where $\ell_b = \lfloor \log_2(b-1)\rfloor$ $\ell_{n-b} = \lfloor \log_2(n-b-1)\rfloor$.

Assume without loss that $\ell_{n-b} \geq \ell_b$. As in Case I, because $\langle n-b, m, P_2\rangle \in W_{\ell_{n-b},P_2}$ and $\langle n, m, P_1\rangle \in W_{k,P_1}$, we must have $\ell_b \leq \ell_{n-b} \leq k-1$. But then we see that there are no values of b that fall into this case at all, since we would have $b \leq 2^{\ell_b+1} \leq 2^k$ and $n-b \leq 2^{\ell_{n-b}} \leq 2^k$ which, by summing, is seen to be in contradiction to the fact that $n \geq 2^{k+1}+1$ from $\langle n, m, P_1\rangle \in W_{k,P_1}$. ■

LEMMA 4.2: *Suppose that* $\langle n, m, P_1\rangle \in U_{k,P_1}$ *for some* k. *Then:*

$$q_{U_k}(n,m) = \max_{b\in[1,n-1]} \left\{1 - \frac{b}{n}q(m,b) - \frac{n-b}{n}q(m,n-b)\right\}, \tag{4.6}$$

and the maximum is achieved at $b = \lfloor \frac{1}{2}n \rfloor$.

PROOF: By symmetry between b and $n-b$, we assume without loss that $1 \leq b \leq \lfloor \frac{1}{2}n \rfloor$. We will show that $1-(b/n)q(m,b) - (n-b/n)q(m,n-b) \leq q_{U_k}(n,m)$ for all possible bids b with equality achieved at $b = \lfloor \frac{1}{2}n \rfloor$. Consider the following cases:

Case I: $\langle b, m, P_2\rangle \in W_{\ell,P_2}$ with $2^\ell < b \leq 2^{\ell+1}$ and $\langle n-b, m, P_2\rangle \in U_{k,P_2}$.

We start with Eq. (4.5) for $1-(b/n)q_{W_\ell}(m,b) - (n-b/n)q_{U_k}(m,n-b)$ and show directly that the difference from $q_{U_k}(n,m)$ is always positive. It is convenient to begin by multiplying out the common denominator:

$$\begin{aligned}
&nm\left(q_{U_k}(n,m) - \left(1 - \frac{b}{n}q_{W_\ell}(m,b) - \frac{n-b}{n}q_{U_k}(m,n-b)\right)\right)\\
&= nm\left(\left(1 - \frac{2^k}{m} + \frac{2}{3}\frac{2^{2k}+2}{mn}\right) - \left(\frac{b}{n}\left(1 - \frac{2^{\ell+1}}{m}\right) + \frac{2^k}{n} + \frac{2}{3}\frac{2^{2\ell+1}-2^{2k}-1}{mn}\right)\right)\\
&= nm - n2^k + \frac{2}{3}\left(2^{2k}+2\right) - b\left(m - 2^{\ell+1}\right) - m2^k - \frac{2}{3}\left(2^{2\ell+1} - 2^{2k} - 1\right)\\
&= n\left(m - 2^k\right) - b\left(m - 2^{\ell+1}\right) - m2^k - \frac{2}{3}2^{2\ell+1} + \frac{4}{3}2^{2k} + 2.
\end{aligned}$$

Now notice that since $\langle n-b, m, P_2\rangle \in U_{k,P_2}$, it must be that $n-b > 2^k$ and $m > 2^k$. Thus, the coefficient of n is positive, so using $n > 2^k + b$ gives:

$$\begin{aligned}
&> \left(2^k + b\right)\left(m - 2^k\right) - b\left(m - 2^{\ell+1}\right) - m2^k - \frac{2}{3}2^{2\ell+1} + \frac{4}{3}2^{2k} + 2\\
&= m2^k + bm - 2^{2k} - b2^k - bm + b2^{\ell+1} - m2^k - \frac{2}{3}2^{2\ell+1} + \frac{4}{3}2^{2k} + 2\\
&= -b\left(2^k - 2^{\ell+1}\right) - \frac{2}{3}2^{2\ell+1} + \frac{1}{3}2^{2k} + 2.
\end{aligned}$$

We notice now that it must be the case that $\ell + 1 \leq k$, because $\langle n-b, m, P_2\rangle \in U_{k,P_2}$ gives $2^{k+1} \geq m$ and $\langle b, m, P_2\rangle \in W_{\ell,P_2}$ gives $m > 2^{\ell+1}$. Hence, the coefficient of b is non-positive,

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

and we can use $b \le 2^{\ell+1}$ to get:

$$\begin{aligned}&\ge -2^{\ell+1}\left(2^k - 2^{\ell+1}\right) - \frac{2}{3}2^{2\ell+1} + \frac{1}{3}2^{2k} + 2\\ &= \frac{2}{3}2^{2(\ell+1)} - 2^{\ell+1}2^k + \frac{1}{3}2^{2k} + 2.\end{aligned}$$

Finally, since $\ell + 1 \le k$, define $\ell + 1 = k - c$, where $c \in \mathbb{N} \cup \{0\}$ and then write the above expression as:

$$= 2^{2k}\left(\frac{2}{3}2^{-2c} - 2^{-c} + \frac{1}{3}\right) + 2.$$

By inspection, $(2/3)2^{-2c} - 2^{-c} + (1/3) = 0$ for $c = 0$ and $c = 1$ and is strictly positive for $c \ge 2$. Thus this expression is always ≥ 0 and we have the desired inequality.

Case II: $\langle b, m, P_2\rangle \in W_{\ell_b,P_2}$ and $\langle n-b, m, P_2\rangle \in W_{\ell_{n-b},P_2}$ with $\ell_b = \lfloor \log_2(b-1)\rfloor$ and $\ell_{n-b} = \lfloor \log_2(n-b-1)\rfloor$.

We have in this case:

$$\begin{aligned}&1 - \frac{b}{n}q_{W_k}(m,b) - \frac{n-b}{n}q_{W_k}(m,n-b)\\ &= 1 - \frac{b}{n}\left(\frac{2^{\ell_b+1}}{m} - \frac{2}{3}\frac{2^{\ell_b+1}-1}{mb}\right) - \frac{n-b}{n}\left(\frac{2^{\ell_{n-b}+1}}{m} - \frac{2}{3}\frac{2^{\ell_{n-b}+1}-1}{m(n-b)}\right)\\ &= 1 + \frac{4}{3nm} - \frac{2}{nm}\left(\left(b - \frac{2}{3}2^{\ell_b}\right)2^{\ell_b} + \left((n-b) - \frac{2}{3}2^{\ell_{n-b}}\right)2^{\ell_{n-b}}\right).\end{aligned}$$

So it satisfies us to minimize the function that appears:

$$\left(b - \frac{2}{3}2^{\ell_b}\right)2^{\ell_b} + \left((n-b) - \frac{2}{3}2^{\ell_{n-b}}\right)2^{\ell_{n-b}}.$$

Since $\ell_b = \lfloor \log_2(b-1)\rfloor$, $\ell_{n-b} = \lfloor \log_2(n-b-1)\rfloor$ we recognize this as $f(b) + f(n-b)$ where $f : \mathbb{N} \to \mathbb{R}$ is the function $f(a) = \left(a - (2/3)2^{\lfloor \log_2(a-1)\rfloor}\right)2^{\lfloor \log_2(a-1)\rfloor}$. But now we observe that this function is convex. Moreover this is a piecewise linear function and can be written as a maximum over a number of linear functions $f(a) = \sup_{i\in\mathbb{N}}\left(a - (2/3)2^i\right)2^i$. Observe then that $f(b) + f(n-b) \ge f\left(\lfloor \frac{1}{2}n\rfloor\right) + f\left(\lceil \frac{1}{2}n\rceil\right)$. In other words, the bid $b = \lfloor \frac{1}{2}n\rfloor$ cannot be beaten. When we plug in $b = \lfloor \frac{1}{2}n\rfloor$ we get exactly the desired bound of $q_{U_k}(n,m)$.

Case III: $\langle b, m, P_2\rangle \in U_{k,P_2}$ and $\langle n-b, m, P_2\rangle \in U_{k,P_2}$

There are no such values of b since we would have $b \ge 2^k + 1$ and $n - b \ge 2^k + 1$ from the definition of U_{k,P_2}, which by summing is seen to contradict $n \le 2^{k+1}$ from the fact that $\langle n, m, P_1\rangle \in U_{k,P_1}$. ■

Remark 4.3: A careful reading of Case II of Lemma 4.2 reveals that for some values of n, there are other bids b that match the performance of $b = \lfloor \frac{1}{2}n\rfloor$. Indeed, the function f is piecewise linear so when $\lfloor \frac{1}{2}n\rfloor$ is on the interior of a linear segment, other bids that have both b and $n - b$ on the same segment will do just as well.

THEOREM 4.4: *(Restatement of Theorem 1.1) We have that $p^\star(n,m) = q(n,m)$ for all n, m and the optimal bidding strategy is $b^\star(n,m) = 2^k$ if $\langle n, m, P_1\rangle \in W_{k,P_1}$ and $b^\star(n,m) = \lfloor \frac{1}{2}n\rfloor$ when $\langle n, m, P_1\rangle \in U_{k,P_1}$.*

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press

PROOF: From Lemma 4.1 and 4.2, we know that q and $p^\star$ both satisfy the same recursion relation from Proposition 2.4. We now prove that $p^\star(n,m) = q(n,m)$ for all (n,m) by induction on the sum $n+m$. For $r \in \mathbb{N}$, $r \geq 3$, let $A_r = \{(n,m) : n+m \leq r\}$ and consider the statement that $p^\star(n,m) = q(n,m)$ for every $(n,m) \in A_r$. The base case $r = 3$ is clear since $p^\star(2,1) = q(2,1) = 0$ and $p^\star(1,2) = q(1,2) = 1$. Now suppose $p^\star(n,m) = q(n,m)$ for every $(n,m) \in A_r$. We have then for any $(n,m) \in A_{r+1}$ that:

$$\begin{aligned} p^\star(n,m) &= \max_{b\in[1,n-1]} \left\{ 1 - \frac{b}{n} p^\star(m,b) - \frac{n-b}{n} p^\star(m,n-b) \right\} \text{ (recurrence for } p^\star) \\ &= \max_{b\in[1,n-1]} \left\{ 1 - \frac{b}{n} q(m,b) - \frac{n-b}{n} q(m,n-b) \right\} \text{ (by induction hypothesis)} \\ &= q(n,m) \text{ (recurrence relation for } q). \end{aligned}$$

The induction hypothesis can be applied since every pair (m,b), $(m,n-b)$ has $m+b \leq r-1$ and $m+n-b \leq r-1$ since $1 \leq b \leq n-1$. By the principle of induction, this shows that $p^\star(n,m) = q(n,m)$ for all $(n,m) \in \cup_{r\in\mathbb{N}} A_r$. Since $\cup_{r\in\mathbb{N}} A_r$ covers the entire quadrant, we have $p^\star(n,m) = q(n,m)$ for all n,m as desired. In light of $p^\star \equiv q$, Lemmas 4.1 and 4.2 also show that $b^\star(n,m) = 2^k$ for $\langle n,m,P_1\rangle \in W_{k,P_1}$ and $b^\star(n,m) = \left\lfloor \frac{1}{2} n \right\rfloor$ when $\langle n,m,P_2\rangle \in U_{k,P_1}$ is the optimal bidding strategy. ■

Acknowledgements

The author was supported by a MacCracken fellowship from New York University and National Science Foundation grant DMS-1209165. The author is very grateful to the anonymous referee for their careful reading of the article and in particular for suggestions that greatly improved the presentation of ideas in Section 3.

References

1. Chen, M.-R. & Hsiau, S.-R. (2010). Two new models for the two-person red-and-black game. *Journal of Applied Probability* 47: 97–108.
2. Condon, A. (1992). The complexity of stochastic games. *Information and Computation* 96: 203–224.
3. Dubins, L.E. & Savage, L.J. (1965). *How to Gamble If You Must: Inequalities for Stochastic Processes*. New York, NY, USA: McGraw-Hill New York, Inc.
4. Eisenberg, B., Stengle, G., & Strang, G. (1993). The asymptotic probability of a tie for first place. *Annals of Applied Probability* 3: 731–745.
5. Haigh, J. & Roters, M. (2000). Optimal strategy in a dice game. *Journal of Applied Probability* 37: 1110–1116.
6. Neller, T. & Presser, C. (2010). Practical play of the dice game pig. *The UMAP Journal* 31: 5–19.
7. Pontiggia, L. (2005). Two-person red-and-black with bet-dependent win probabilities. *Advances in Applied Probability* 37: 75–89.
8. Secchi, P. (1997). Two-person red-and-black stochastic games. *Journals of Applied Probability* 34: 107–126.
9. Shapley, L.S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences of the United States of America* 39: 1095–1100.
10. Tijms, H. & van der Wal, J. (2006). A real-world stochastic two-person game. *Probability in the Engineering and Informational Sciences* 20: 599–608.
11. van de Brug, T., Kager, W., & Meester, R. (2015). The Asymptotics of Group Russian Roulette. arXiv 1507.03805. Preprint. http://arxiv.org/abs/1507.03805
12. Wikipedia (2015). Guess Who? — Wikipedia, the free encyclopedia. http://en.wikipedia.org/w/index.php?title=Guess%20Who%3F

https://doi.org/10.1017/S026996481600022X Published online by Cambridge University Press