European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Optimal strategies in the fighting fantasy gaming system: Influencing stochastic dynamics by gambling with limited resource

Iain G. Johnston[a,b]

[a] Department of Mathematics, Faculty of Mathematics and Natural Sciences, University of Bergen, Bergen, Norway
[b] Alan Turing Institute, London, UK

ARTICLE INFO

Article history:
Received 13 August 2019
Accepted 24 January 2022
Available online xxx

ABSTRACT

In many games and other processes, participants can choose to intervene in some way that does not follow the usual progress of the game (for example, cheating at cards, or spying on rivals) which may provide benefits, but also possibly incur substantial costs. Here, repeated interventions may be more likely to incur negative outcomes – for example, as the chance of getting caught increases. How to optimally employ these risky interventions, trading off potential advantages and disadvantages, can then be challenging to identify. Here, we study such a game, taken from the popular 'Fighting Fantasy' gamebook series. This stochastic game involves a series of rounds, each of which may be won or lost. Each round, a unit of limited resource ('LUCK') may be spent on a gamble to amplify benefits from a win or to mitigate deficits from a loss. However, the success of this gamble depends on the number of units of remaining resource, and if the gamble is unsuccessful, benefits are reduced and deficits increased. By choosing to expending resource, a player thus has diminishing probability of positive return, as in the cheating and espionage examples above. We characterise the dynamics of this system using stochastic analysis and dynamic programming, solve the Bellman equation for the complete system with diminishing returns, and identify the optimal strategy for any given state during the game. We use classification tools to distil general principles for this and related problems, informing resource allocation problems with diminishing returns in stochastic decision theory.

© 2022 The Author(s). Published by Elsevier B.V.
This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

1. Introduction

Many processes have the potential for a participant to intervene in an attempt to influence the outcome of a particular aspect of play. In instances like cheating, such interventions come at a cost. The player may gain some advantages from intervening, but if they are discovered, disadvantages are more likely. The question of whether, and how often, to intervene in this way must therefore trade off the potential advantages from intervening against the possibility of these negative outcomes. In particular, if repeated interventions are more likely to incur a negative outcome, the optimal intervention strategy may be challenging to identify.

Here, we study a stochastic decision problem with this structure, taken from the 'Fighting Fantasy' (FF) adventure gaming system, where a single player takes part in an interactive fiction story (Costikyan, 2007; Green, 2014; Jackson & Livingstone, 2002). Here, the reader makes choices that influence the progress through the book, the encounters that occur, and the outcomes of combats. In many cases, die rolls are used to provide stochastic influence over the outcomes of events in these games, particularly combat dynamics. These combat dynamics affect the game outcome (and thus the experience of millions of players worldwide) yet have rarely been studied in detail.

In FF, a player is assigned statistics (SKILL and STAMINA), dictating combat proficiency and endurance respectively. Opponents are also characterised by these combat statistics. Combat proceeds iteratively through a series of 'attack rounds'. In a given round, according to die rolls, the player may draw, win or lose, respectively. These outcomes respectively have no effect, damage the opponent, and damage the player. The player then has the option of using a limited resource (LUCK) to apply control to the outcome of the round. This decision can be made dynamically, allowing the player to choose a policy based on the current state of the system. However, each use of LUCK is a gamble (Dubins & Savage, 1965; Maitra & Sudderth, 2012), where the probability of success depends on the current level of the resource. If this gamble is successful, the player experiences a positive outcome (damage to the opponent is

E-mail address: iain.johnston@uib.no

https://doi.org/10.1016/j.ejor.2022.01.039
0377-2217/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/)

amplified; damage to the player is weakened). If the gamble is unsuccessful, the player experiences a negative outcome (damage to the opponent is weakened, damage to the player is amplified). The optimal strategy for applying this control in a given state has yet to be found.

The concepts of SKILL, STAMINA, and LUCK can be thought of more generally outside the FF context. SKILL determines an agent's prowess at the game; this may be a player's skill at cards, or a firm's propensity to develop leading products. STAMINA determines the number of losses a player can absorb without losing outright; the number of tricks in a card game, or the number of loss-making products that can be produced before bankruptcy. Depending on context, LUCK could correspond to properties of a player including trustworthiness or concealment; broadly, reflecting the probability that an intervention will be successful. This is the resource that is diminished as interventions are employed. The question then becomes, against an opponent of known SKILL and STAMINA, when and how should a player choose to intervene using LUCK to increase the probability of overall victory?

This is a stochastic decision problem on a discrete state space. The system is Markovian: in the absence of special rules, the statistics of the player and opponent uniquely determine a system state, and this state combined with a choice of policy uniquely determine the transition probabilities to the next state. The problem of determining the optimal strategy is then a Markov decision problem (MDP) (Bellman, 1957; Kallenberg, 2003). In an MDP, a decision-maker must choose a particular strategy for any given state of a system, which evolves according to Markovian dynamics. In FF combat, the decision is always binary: given a state, whether or not to use a unit of the diminishing resource of LUCK to attempt to influence the outcome of a given round.

The study of stochastic games and puzzles is long established in operational research (Bellman, 1965; Smith, 2007) and has led to several valuable and transferrable insights (Little, Murty, Sweeney, & Karel, 1963; Smith, 2007). Markov analysis, dynamic programming, and simulation have been recently used to explore strategies and outcomes in a variety of games, sports, and TV challenges (Clarke & Norman, 2003; Johnston, 2016; Lee, 2012; Percy, 2015; Perea & Puerto, 2007; Smith, 2007). Specific analyses of popular recreational games with a stochastic element including Solitaire (Kuykendall & Mackenzie, 1999; Rabb, 1988), Flip (Trick, 2001), Farmer Klaus and the Mouse (Campbell, 2002), Tetris (Kostreva & Hartman, 2003), and The Weakest Link (Thomas, 2003). These approaches typically aim to identify the optimal strategy for a given state, and, in win/lose games, the overall probability of victory over all possible instances of the game (Smith, 2007). In stochastic dynamic games, counterintuitive optimal strategies can be revealed through mathematical analysis, not least because 'risking points is not the same as risking the probability of winning' (Neller & Presser, 2004).

The FF system has some conceptual similarities with the well-studied recreational game Pig, and other so-called 'jeopardy race games' (Neller & Presser, 2004; Smith, 2007), where die rolls are used to build a score then a decision is made, based on the current state of the system, whether to gamble further or not. Neller & Presser have used a value iteration approach to identify optimal strategies in Pig and surveyed other similar games (Neller & Presser, 2004). In FF combat, however, the player has potential agency both over their effect on the opponent and the opponent's effect on them. Further, resource allocation in FF is a dynamic choice and also a gamble (Dubins & Savage, 1965; Maitra & Sudderth, 2012), the success probability of which diminishes as more resource is allocated. The probability of a negative outcome, as opposed to a positive one, therefore increases as more resource is used, providing an important 'diminishing returns' consideration in policy decision (Deckro & Hebert, 2003). In an applied context this could correspond to engaging in, for example, espionage and counterespionage (Solan & Yariv, 2004), with increasing probability of negative outcomes with more engagement in these covert activites.

The optimal policy for allocating resource to improve a final success probability has been well studied in the context of research and development (R&D) management (Baye & Hoppe, 2003; Canbolat, Golany, Mund, & Rothblum, 2012; Gerchak & Parlar, 1999; Heidenberger & Stummer, 1999). While policies in this field are often described as 'static', where an initial 'up-front' decision is made and not updated over time, dynamic policy choices allowing updated decisions to be made based on the state of the system (including the progress of competitors) have also been examined (Blanning, 1981; Hopp, 1987; Posner & Zuckerman, 1990). Rent-seeking 'contest' models (Clark & Riis, 1998) also describe properties of the victory probability as a function of an initial outlay from players. The 'winner takes all' R&D model of Canbolat et al., where the first player to complete development receives all the available payoff, and players allocate resource towards this goal (Canbolat et al., 2012), bears some similarity to the outcomes of the FF system. The model of Canbolat et al. did not allow dynamic allocation based on the current system state, but did allow a fixed cost to be spread over a time horizon, and computed Nash equilibria in a variety of cases under this model.

A connected branch of the literature considers how to allocate scarce resource to achieve an optimal defensive outcome (Golany, Goldberg, & Rothblum, 2015; Valenzuela, Szidarovszky, & Rozenblit, 2015), a pertinent question both for human (Golany, Kaplan, Marmur, & Rothblum, 2009) and animal (Clark & Harvell, 1992) societies. Inspection games, a non-cooperative situation where limited resource is used by competing parties to verify, or display, adherence to legal rules are a particular pertinent example (Deutsch, 2021; Deutsch, Golany, & Rothblum, 2011; Dresher, 1962). Both optimisation and Nash equilibrium approaches are used in these contexts to identify solutions to the resource allocation problem under different structures (Golany et al., 2015; Valenzuela et al., 2015). The FF system has such a defensive component, but the same resource can also be employed offensively, and as above takes the less-studied form of a gamble with a diminishing success probability. We will attempt to analyse the dynamics of this system and show how its behaviour can inform the broader class of decision problems where possible interventions are gambles that costs resource and decrease future success probabilities.

1.1. Game dynamics

Within an FF game, the player has nonnegative integer statistics called SKILL, STAMINA, and LUCK. SKILL and LUCK are typically ≤ 12; STAMINA is typically ≤ 24, although these bounds are not required by our analysis. Combats, the focus of this study, occur throughout an FF game, as the player encounters different adversaries within the unfolding interactive fiction storyline. In a given combat, the opponent will also have SKILL and STAMINA statistics. We label the SKILL, STAMINA, and LUCK of the player (the 'hero') as k_h, s_h, and l respectively, and the opponent's SKILL and STAMINA as k_o and s_o. Broadly, combat in the FF system involves a series of rounds, where differences in SKILL between combatants influences how much STAMINA is lost in each round; when one combatant's STAMINA reaches zero or below, the combat is over and that combatant has lost. The player may choose to use LUCK in any given round to influence the outcome of that round. More specifically, combat proceeds through Algorithm 1.

Algorithm 1. Fighting Fantasy combat system Jackson & Livingstone (2002). The game takes place between a player 'hero' with SKILL k_h, initial STAMINA s_h, and initial LUCK l, and an opponent with SKILL k_o and initial STAMINA s_o.

1. Roll two dice and add k_h; this is the player's attack strength A_h.
2. Roll two dice and add k_o; this is the opponent's attack strength A_o.
3. If $A_h = A_o$, this attack round is a draw. Go to 6.
4. If $A_h > A_o$, the player has won this attack round. *Make decision* whether to use LUCK.
   (a) If *yes*, roll two dice to obtain r. If $r \le l$, set $s_o = s_o - 4$. If $r > l$, set $s_o = s_o - 1$. For either outcome, set $l = l - 1$. Go to 6.
   (b) If *no*, set $s_o = s_o - 2$. Go to 6.
5. If $A_h < A_o$, the opponent has won this attack round. *Make decision* whether to use LUCK.
   (a) If *yes*, roll two dice to obtain r. If $r \le l$, set $s_h = s_h - 1$. If $r > l$, set $s_h = s_h - 3$. For either outcome, set $l = l - 1$. Go to 6.
   (b) If *no*, set $s_h = s_h - 2$. Go to 6.
6. If $s_h > 0$ and $s_o \le 0$, the player has won; if $s_o > 0$ and $s_h \le 0$, the opponent has won. Otherwise go to 1.

2. Analysis

In basic combat dynamics, SKILL does not change throughout a combat. The probabilities p_w, p_d, p_l of winning, drawing, an losing an attack round depend only on SKILL, and remain constant throughout the battle. We therefore consider $P^t(s_h, s_o, l)$: the probability, at round t, of being in a state where the player has STAMINA s_h and LUCK l, and the opponent has STAMINA s_o.

We are concerned with the *victory probability* v_p with which the player is eventually victorious, corresponding to a state where $s_h \ge 1$ and $s_o \le 0$. We thus consider the 'getting to a set' outcome class of this stochastic game (Maitra & Sudderth, 2012), corresponding to a 'winner takes all' race (Canbolat et al., 2012). We start with the probability of winning, drawing, or losing a given round. First, let k_h be the player's SKILL and k_o be the opponent's SKILL. The important associated quantity in determining the system's behaviour is the difference in SKILL $\Delta_k = k_h - k_o$.

The outcome of each round of the game is determined by a comparison between $k_h + X_1$ and $k_o + X_2$, where X_1, X_2 are random variates describing the sum of two dice (steps 3–5 in Algorithm 1). Specifically, if $k_h + X_1 > k_o + X_2$ the player wins the round, if $k_h + X_1 < k_o + X_2$ the player loses the round, and if the two are equal the round is drawn. Rearranging, we obtain the inequalities $X_1 - X_2 + \Delta_k > 0$, $X_1 - X_2 + \Delta_k < 0$, and $X_1 + X_2 + \Delta_k = 0$ for win, loss, and draw respectively. If we define the new random variable $X = X_1 - X_2$ (hence, the difference between the sums of two sets of two dice), the probabilities of win, draw, and loss events correspond respectively to $p_w = P(X + \Delta_k > 0)$, $p_d = P(X + \Delta_k = 0)$, $p_l = P(X + \Delta_k < 0)$.

For our dice-based system, X follows a discrete distribution on $[-10, 10]$. The point density $P(X = i)$ is $\frac{1}{1296}\{1, 4, 10, 20, 35, 56, 80, 104, 125, 140, 146, 140, 125, 104, 80, 56, 35, 20, 10, 4, 1\}$; (OEIS sequence A063260 (OEIS, 2019)). Hence the probabilities of a round being won, lost, or drawn can be computed by summing probabilities over the values of X that are compatible with each outcome:

$$p_w(\Delta_k) = \sum_{j=-\Delta_k+1}^{10} P(X = j) \tag{1}$$

$$p_d(\Delta_k) = P(X = -\Delta_k) \tag{2}$$

$$p_l(\Delta_k) = \sum_{j=-10}^{-\Delta_k-1} P(X = j) \tag{3}$$

for FF (although our analysis can be applied to any p_w, p_d, p_l to suit other model situations).

2.1. Dynamics without interventions

We first consider the straightforward case where the player employs no strategy, never electing to use LUCK. This case corresponds to the 'default' dynamics of the game, without any interventions from the player – following the broader examples above, this corresponds to no use of cheating or espionage. We can then ignore l and consider steps through the (s_h, s_o) STAMINA space, which form a discrete-time Markov chain. The probability of a given state in the next timestep is then

$$P^{t+1}(s_h, s_o, l) = p_w P^t(s_h, s_o + 2, l) + p_l P^t(s_h + 2, s_o, l) + p_d P^t(s_h, s_o, l). \tag{4}$$

We can consider a combinatorial approach based on 'game histories' describing steps moving through this space (Maitra & Sudderth, 2012). Here a game history is a string from the alphabet $\{W, D, L\}$, with the character at a given position i corresponding to respectively to a win, draw, loss in round i. We aim to enumerate the number of possible game histories that correspond to a given outcome, and assign each a probability.

We write w, d, l for the character counts of W, D, L in a given game history. A victorious game must always end in W. Consider the string describing a game history omitting this final W. First leaving out Ds, we have $(w-1)$ Ws and l Ls that can be arranged in any order. We therefore have $n(w, l) = \binom{w-1+l}{l}$ possible strings, each of length $w - 1 + l$. For completeness, we can then place any number d of Ds within these strings, obtaining

$$n(w, l, d) = \binom{w-1+l}{l}\binom{w-1+l+d}{d}. \tag{5}$$

Write $\sigma_h = \lceil s_h/2 \rceil$ and $\sigma_o = \lceil s_o/2 \rceil$, describing the number of rounds each character can lose before dying. Then, for a player victory, $w = \sigma_o$ and $l \le \sigma_h - 1$. d can take any nonnegative integer value. The appearance of each character in a game string is accompanied by a multiplicative factor of the corresponding probability, so we obtain

$$v_p = p_w^{\sigma_o} \sum_{l=0}^{\sigma_h-1} \sum_{d=0}^{\infty} p_l^l p_d^d \binom{w-1+l}{l}\binom{w-1+l+d}{d}, \tag{6}$$

where the probability associated with the final W character has now also been included. Using hypergeometric functions, a closed form for this expression can be written down (see Supplementary Information), but as p_d is small, the sum over d converges after a small number of terms and Eq. (6) can readily be computed numerically.

Equation (6) is compared with the result of stochastic simulation in Fig. 1, and shown for various SKILL differences Δ_k and (s_h, s_o) initial conditions. Here, the stochastic simulation simply involves 10^3 simulations of Algorithm 1, with the given initial conditions, and with uniform random variates used for die rolls (code for this simulation and other analyses throughout are available at https://github.com/StochasticBiology/fighting-fantasy-analysis). Intuitively, more favourable $\Delta_k > 0$ increase v_p and less favourable $\Delta_k < 0$ decrease v_p for any given state, and discrepancies between starting s_h and s_o also influence eventual v_p. A pronounced s (mod 2) structure is observed, as in the absence of LUCK, $s = 2n$ is functionally equivalent to $s = 2n - 1$ for integer n. For lower initial STAMINAS, v_p distributions become more sharply peaked with s values, as fewer events are required for an eventual outcome. The general interpretation of these 'default' dynamics is, intuitively, that victory probability in the absence of interventions is higher for the combatant with the higher probability of winning individual rounds, and the ability to absorb more round losses. One of these factors may compensate for the other; for example, if the player has a SKILL disadvantage of $\Delta_k = -2$ but has three times their opponents' STAMINA, victory is still likely. Different player styles –

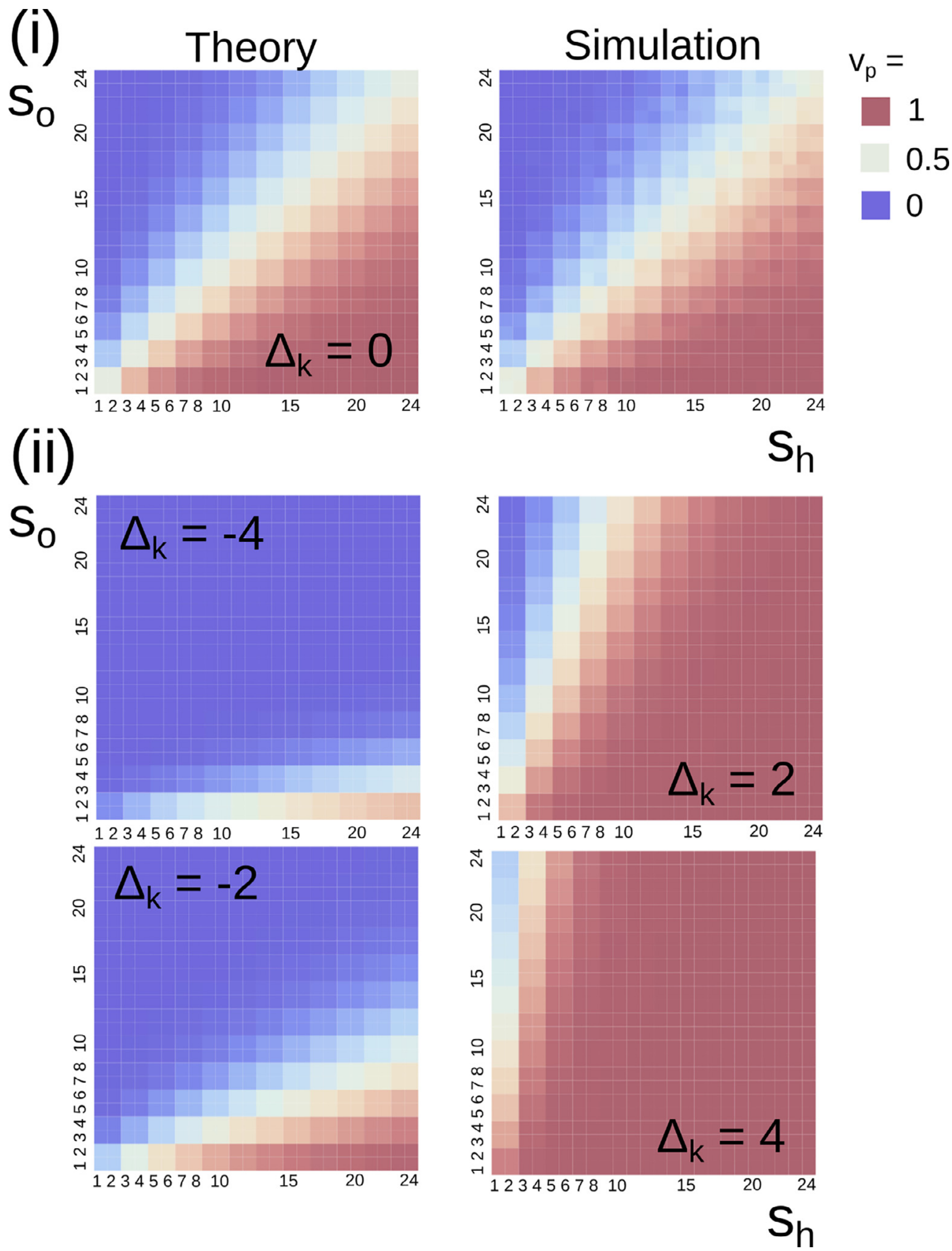

Fig. 1. Victory probability in the absence of LUCK -based strategy. (i) Comparison of predicted victory probability v_p from Eq. (6) with stochastic simulation. (ii) v_p behaviour as SKILL difference Δ_k changes.

skillful but susceptible to damage, or loss-prone but robust – may thus be equally successful.

2.2. Analytic dynamics with interventions but no diminishing resource

We next consider a simplified case where the player can intervene in an attempt to alter the 'default' dynamics of the game. Specifically, to increase the probability of victory beyond the basic case in Eq. (6), the player can elect to use LUCK in any given round. We will first demonstrate that the above history-counting analysis can obtain analytic results when applied to a simplified situation when LUCK is not depleted by use, so that the only limit on its employment is its initial level (Blanning, 1981). This corresponds more broadly to the case where the probability of an adverse outcome does not change with the number of interventions that are used – so, for example, the probability of getting caught cheating remains the same no matter how many times the player cheats.

First consider the case where LUCK is only used to amplify positive outcomes of individual rounds (step 4a in Algorithm 1), and is used in every successful round. For now, ignore losses and draws. Then every game history consists of *A*s and *B*s, where *A* is a successful offensive use of LUCK and *B* is an unsuccessful offensive use of LUCK. Consider the game histories that lead to the opponent losing exactly n STAMINA points *before* the final victorious round. There are $\lfloor n/4 \rfloor + 1$ string lengths that can achieve this, which are $L = n - 3k$, where k runs from 0 to $\lfloor n/3 \rfloor$. The strings with a given k involve $n - 4k$ failures and k successes.

If we make the simplifying assumption that LUCK is not depleted with use, every outcome of a LUCK test has the same success probability $q(l) = q$. Then the problem is simplified to finding the number of ways of arranging k As and $(n-4k)$ Bs for each possible string:

$$N(k) = \frac{(n-3k)!}{k!(n-4k)!} \tag{7}$$

Now, for every string with a given k, with corresponding string length $n-3k$, we can place l Ls and d Ds as before, giving

$$N(k; n, l, d) = \frac{(n-3k)!}{k!(n-4k)!}\binom{n-3k+l}{l}\binom{n-3k+l+d}{d}. \tag{8}$$

The complete history involves a final victorious round. For now we will write the probability of this event as p_f, then the probability associated with this set of histories is

$$P(k; n, l, d, p_f) = p_f \frac{(n-3k)!}{k!(n-4k)!} q^k (1-q)^{n-4k} p_l^l \binom{n-3k+l}{l} p_d^d \binom{n-3k+l+d}{d} \tag{9}$$

with

$$P(k; n, p_f) = \sum_{l=0}^{\sigma_h - 1}\sum_{d=0}^{\infty} P(k; n, l, d, p_f) \tag{10}$$

Again, a closed-form solution for this expression exists using hypergeometric functions (see Supplementary Information) but the sums converge quickly and so can readily be computed numerically. Finally

$$P_n(n, p_f) = \sum_{k=0}^{n/3} P(k; n, p_f). \tag{11}$$

Now consider the different forms that the final victorious round can take. The opponent's STAMINA can be reduced to 4 followed by an A, 3 followed by A, 2 followed by A, or 1 followed by A or B. If we write $P(m; X)$ for the probability of reducing the opponent's STAMINA to m then finishing with event X,

$$v_p = P(4; A) + P(3; A) + P(2; A) + P(1; A) + P(1; B) \tag{12}$$

hence

$$v_p = P_n(s_o - 4, q) + P_n(s_o - 3, q) + P_n(s_o - 2, q) + P_n(s_o - 1, q) + P_n(s_o - 1, (1-q)). \tag{13}$$

Similar expressions can be derived for the defensive case, where LUCK is solely used when a round is lost, and with some relaxations on the structure of the sums involved the case where LUCK is not used in every round can also be considered. Fig. 2(i) compares Eq. (13) and stochastic simulation, and shows that use of LUCK can dramatically increase victory probability in a range of circumstances. For example (Fig. 2(ii)), high LUCK used judiciously can almost compensate for a $\Delta_k = -2$ SKILL disadvantage, and make victory highly likely if player and opponent have the same SKILL (the only exception being when opponent STAMINA greatly exceeds player STAMINA). Notably, even when the success of an intervention is far from certain ($l = 7$), substantial shifts in victory probability are possible. Phrased more generally, this simpler case demonstrates the dramatic positive shifts in game outcome possible from interventions, albeit when the diminishing nature of the intervention resource is not accounted for.

2.3. Stochastic optimal control with dynamic programming for full system with limited and diminishing resource

We now consider the complete system, where electing to intervene diminishes the resource that determines the success of interventions. Each decision to use LUCK now diminishes LUCK by one. We can no longer use the simple counting argument in Eq. (9) to compute the probabilities of each history, because each probability now depends on the specific structure of the history. It will be possible to enumerate these histories exhaustively but the analysis rapidly expands beyond the point of useful interpretation, so we turn to dynamic programming to investigate the system's behaviour.

In a game with a given Δ_k, we characterise every state of the system with a tuple $\mathcal{S} = \{s_h, s_o, l, O\}$ where O is the outcome (win or loss) of the current attack round. The question is, given a state, should the player elect to intervene (use LUCK) or not?

A common approach to identify the optimal strategy for a Markov decision problem in a discrete state space is to use the Bellman equation (Bellman, 1957; Kirk, 2012), which in our case is simply

$$v_p(\mathcal{S}) = \max_a \left(\sum_{\mathcal{S}'} P_a(\mathcal{S}, \mathcal{S}') v_p(\mathcal{S}') \right) \tag{14}$$

$$v_p(\mathcal{S}) = 0, \text{ for all } \mathcal{S} \text{ with } s_h \le 0 \text{ and } s_o > 0 \tag{15}$$

$$v_p(\mathcal{S}) = 1, \text{ for all } \mathcal{S} \text{ with } s_h > 0 \text{ and } s_o \le 0 \tag{16}$$

Here, a is a strategy dictating what action to take in state $\mathcal{S}$, $P_a(\mathcal{S}, \mathcal{S}')$ is the probability under strategy a of the transition from state $\mathcal{S}$ to state $\mathcal{S}'$, and $v_p(\mathcal{S})$ is the probability-to-victory of state $\mathcal{S}$. The joint problem is to compute the optimal v_p, and the strategy a that maximises it, for all states. To do so, we employ a dynamic programming approach of backward induction (Bellman, 1957), starting from states where v_p is known and computing backwards through potential precursor states.

The dynamic programming approach first assigns a probability-to-victory v_p for those states of the system where the outcome is immediately defined (Supplementary Fig. S1(i)). Hence, for all 'defeat' states with $s_h \le 0, s_o > 0$, we set $v_p = 0$; for all 'victory' states with $s_o \le 0, s_h > 0$, we set $v_p = 1$. States where both $s_h \le 0$ and $s_o \le 0$ are inaccessible under the dynamics of the system and are thus ignored. We then iteratively consider all states in the system where an outcome leads to a state where v_p is fully determined. For example, after defining $v_p = 0$ for all 'defeat' states, we can compute v_p for all states where a loss outcome leads to a 'defeat' state (Supplementary Fig. S1(ii)). Then we can compute v_p for all states that lead to those states, and so on. We can proceed similarly for states where a win outcome leads to a 'victory' states (Supplementary Fig. S1(iii)), then states leading to those states, and so on.

For states involving a loss outcome, we compute two probability propagators. The first corresponds to the strategy where the player elects to use LUCK, and is of magnitude

$$\begin{aligned} p_y &= \sum_{\mathcal{S}'} P_y(\mathcal{S}, \mathcal{S}') v_p(\mathcal{S}') \\ &= q(l) p_l v_p(\{s_h - 1, s_o, l-1, 0\}) \\ &\quad + (1 - q(l)) p_l v_p(\{s_h - 3, s_o, l-1, 0\}) \\ &\quad + q(l) p_w v_p(\{s_h - 1, s_o, l-1, 1\}) \\ &\quad + (1 - q(l)) p_w v_p(\{s_h - 3, s_o, l-1, 1\}) \end{aligned} \tag{17}$$

the second corresponds to the strategy where the player does not use LUCK, and is

$$\begin{aligned} p_n &= \sum_{\mathcal{S}'} P_n(\mathcal{S}, \mathcal{S}') v_p(\mathcal{S}') \\ &= p_l v_p(\{s_h - 2, s_o, l, 0\}) + p_w v_p(\{s_h - 2, s_o, l, 1\}). \end{aligned} \tag{18}$$

When considering the probability-to-victory for a given state, we hence consider both the next (s_h, s_o, l) combination that a

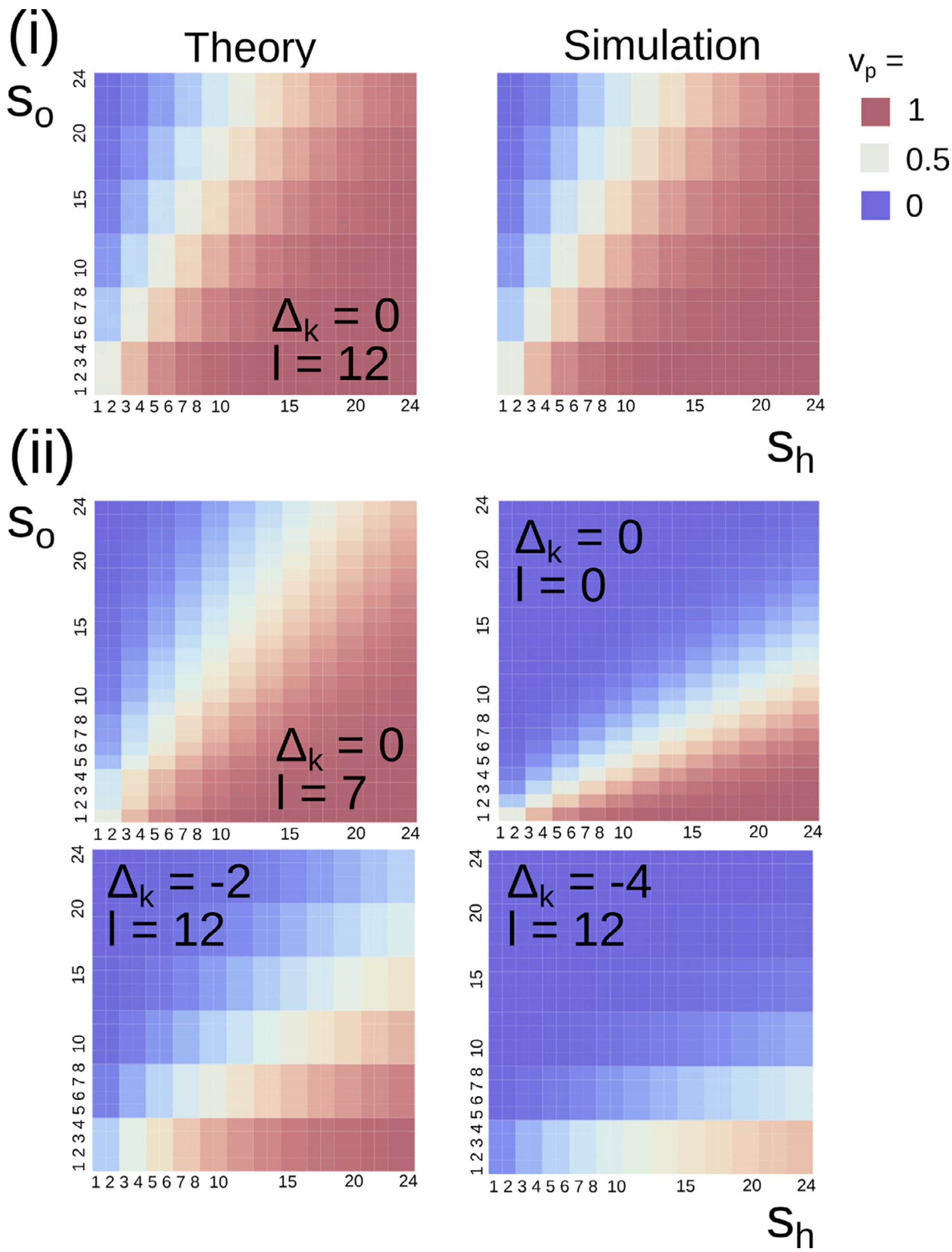

Fig. 2. Victory probability with constant LUCK employed offensively. (i) Eq. (13) compared to stochastic simulation. (ii) Constant luck outcomes for intermediate and low LUCK, and for high LUCK mitigating low Δ_k values.

given event will lead to, and also both possible outcomes (win or loss) from this state. For a given state, if $p_y > p_n$, we record the optimal strategy as using LUCK and record $v_p = p_y$; otherwise we record the optimal strategy as not to use LUCK and record $v_p = p_n$. In practise we replace $p_y > p_n$ with the condition $p_y > (1 + 10^{-10})p_n$ to avoid numerical artefacts, thus requiring that the use of LUCK has a *relative* advantage to v_p above 10^{-10}.

We do the same for states involving a win outcome from this round, where the two probability propagators are now

$$\begin{aligned} p_y = {} & q(l) p_l v_p(\{s_h, s_o - 4, l - 1, 0\}) \\ & + (1 - q(l)) p_l v_p(\{s_h, s_o - 1, l - 1, 0\}) \\ & + q(l) p_w v_p(\{s_h, s_o - 4, l - 1, 1\}) \\ & + (1 - q(l)) p_w v_p(\{s_h, s_o - 1, l - 1, 1\}); \end{aligned} \tag{19}$$

and

$$p_n = p_l v_p(\{s_h, s_o - 2, l, 0\}) + p_w v_p(\{s_h, s_o - 2, l, 1\}). \tag{20}$$

Each new pair of states for which the optimal v_p is calculated opens up the opportunity to compute v_p for new pairs of states (Supplementary Fig. S1(iv)). Eventually a v_p and optimal strategy is computed for each outcome, providing a full 'roadmap' of the optimal decision to make under any circumstance. This full map is shown in Supplementary Fig. S2, with a subset of states shown in Fig. 3.

We find that a high LUCK score and judicious use of LUCK can dramatically enhance victory probability against some opponents. As an extreme example, with a SKILL detriment of $\Delta_k = -9$, $s_h = 2, s_o = 23, l = 12$, use of LUCK increased victory probability by a factor of 10^{18}, albeit to a mere 2.3×10^{-19}. On more reasonable scales, with a SKILL detriment of $\Delta_k = -4$, $s_h = 22, s_o =$

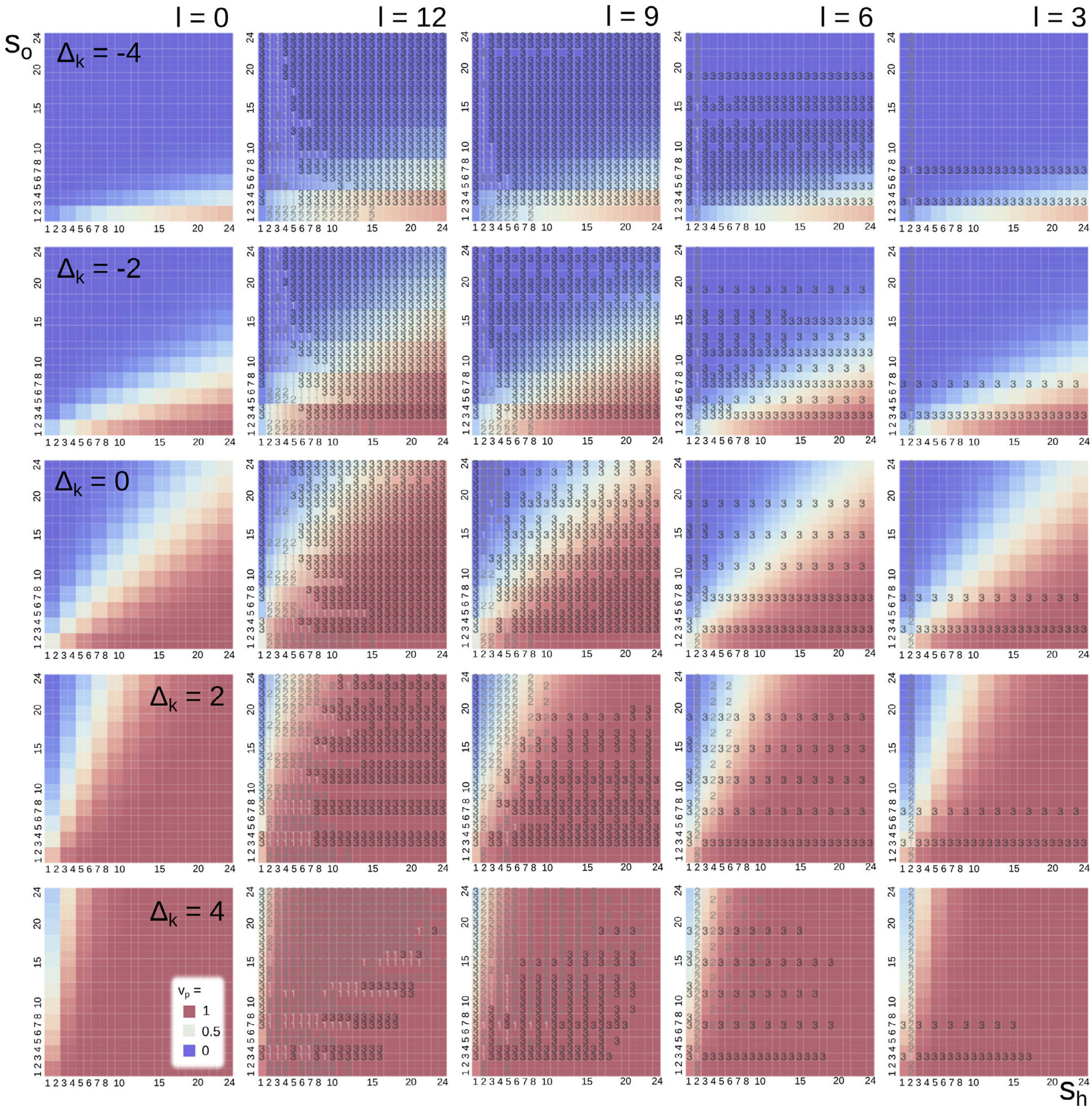

Fig. 3. Optimal strategies and victory probabilities throughout state space. The optimal strategy at each state is given by the numerical code at each corresponding point in the figure: (1 – use LUCK regardless of the outcome of this round; 2 – use LUCK if this round is lost; 3 – use LUCK if this round is won). No number means that the optimal strategy is not to employ LUCK regardless of outcome. Colour gives victory probability v_p.

19, $l = 12$, use of LUCK increased victory probability 1159-fold, from 9.8×10^{-6} to 0.011 (the highest fold increase with the final probability greater than 0.01). With a SKILL detriment of $\Delta_k = -2$, $s_h = 22, s_o = 21, l = 12$, use of LUCK increased victory probability 21-fold, from 0.010 to 0.22 (the highest fold increase with the initial probability greater than 0.01). Using encounters that appear in the FF universe (Gascoigne, 1985), for a player with maximum statistics $k_h = 12, s_h = 24, l = 12$, optimal use of LUCK makes victory against an adult White Dragon ($k_o = 15, s_o = 22$) merely quite unlikely ($v_p = 0.046$) rather than implausible ($v_p = 4.4 \times 10^{-4}$), and victory against a Hell Demon ($k_o = 14, s_o = 12$) fairly straightforward ($v_p = 0.78$) rather than unlikely ($v_p = 0.28$). Outside of the specific bounds of FF, it is clear that judicious choice of intervention strategy, even with limited and diminishing resource, can substantially improve game outcomes, supporting fold changes and even order-of-magnitude increases in victory probability.

2.4. Structure of optimal policy space

We first report some specific features of the optimal policy space characterised in Supplementary Fig. S2, before describing some more general principles of this (and other) systems. There is substantial similarity in optimal policy choice between several regions of state space. For large SKILL deficiencies $\Delta_k \leq -6$ (low probability of victory) the distribution of optimal strategies in STAMINA space is the same for a given l for all Δ_k. For higher l, this similarity continues to higher Δ_k; for $l = 12$, only 5 points in STAMINA space have different optimal strategies for $\Delta_k = -9$ and $\Delta_k = -4$. At more reasonable victory probabilities, a moderate transition is apparent between $l = 6$ and $l = 5$, where the number of points in STAMINA space where the optimal strategy involves using LUCK decreases noticeably (reflecting the lower expected advantage for $l = 5$).

The interplay of several general strategies is observed in the optimal structures. First note that $\lceil s/4 \rceil$ gives the number of hits required for defeat if a hit takes 4 STAMINA points (a successful offensive LUCK test) and $\lceil s/2 \rceil$ gives the number of hits required for defeat in the absence of strategy. These scales partition STAMINA space by the number of rounds required for a given outcome and hence dictate several of the 'banded' structures observable in strategy structure. For example, at $s_h = 2$, it is very advantageous for the player to attempt to mitigate the effect of losing another round. Almost all circumstances display a band of defensive optimal strategy at $s_h = 2$.

At $s_o = 3$, a successful offensive LUCK test is very advantageous (immediate victory). An unsuccessful offensive test, leading to $s_o = 2$, is not disadvantageous to the same extent: we still need exactly one successful attack round without LUCK, as we would if we had not used LUCK and achieved $s_o = 1$ instead. A strip of offensive strategy (or joint offensive-defensive strategy) at $s_o = 3$ is thus the next most robustly observed feature, disappearing only when victory probability is already overwhelmingly high. Many other structural features result from a tradeoff between conserving LUCK and increasing the probability of encountering this advantageous region. An illustration of the broad layout of optimal strategies is shown in Supplementary Fig. S3; a more fine-grained analysis is provided in the Supplementary Information.

While the dynamic programming approach above gives the optimal strategy for any circumstance, the detailed information involved does not lend itself to easy memorisation. As in Smith's discussion of solitaire, 'the curse of dimensionality applies for the state description, and most wise players use heuristics' (Smith, 2007). We therefore consider, in addition to the semi-quantitative summary in Supplementary Fig. S3, coarse-grained quantitative 'strategies' that, rather than specifying an action for each branch of the possible tree of circumstances, use heuristic rules that can be applied in all circumstances. To this end, we computed a classification tree for the optimal strategy given current state variables Δ_k, l, s_h, s_o (Fig. 4), using the *rpart* library in R Team (2018). This decision tree bases an outcome on various inequalities describing the current state of the system, and determines 58% of the optimal policy structure correctly, while dramatically coarse-graining the 'strategy map'. The tree functions as a dynamic decision guide, to be consulted in each round of the game, not just at the initial state.

Several general insights can also be determined from the structure of Fig. 4(i). In determining whether or not to intervene, the first-level questions are about the amount of resource (LUCK l) available to do so. If $l < 6.5$, a negative intervention outcome is more likely than a positive one, and interventions are classified as reserved for unusual circumstances. If $l > 6.5$, a positive outcome is more likely, and intervention is more broadly favoured. In both cases, the next-level question is about the propensity to win a given round (Δ_k). This determines the final levels of questions, which are to do with the specific amount of losses and wins possible before victory or defeat (s_h and s_o).

Specifically, when negative outcomes are more likely, there is an 'emergency' pathway where resource is used defensively to avert the final loss leading to defeat and an 'overwhelmed' pathway where resource is used offensively when the opponent can absorb many loss events. When positive outcomes are likely, the defensive use of resource is favoured unless the player can absorb many losses, and offensive use generally favoured to improve victory probability.

The specific inequalities involved in this coarse-grained decision structure depend on the FF system, but the principles they embody can be generalised to systems that quantitatively differ. To verify this, we adapted the system to reflect a quantitatively different case, where a successful intervention after a won round (step 4a in Algorithm 1) causes 3 damage to the opponent (not the usual 4). This adaptation changes the structure of optimal policy space (Supplementary Fig. S4)) but the resulting classification tree (now identifying 45% of the optimal choices in the spaces) displays the same coarse-grained strategies (Supplementary Fig. S5; Fig. 4(ii)).

3. Discussion

We have examined the probability of victory in an iterated, probabilistic decision problem that plays a central role in a well-known and widespread interactive fiction series. The game can be played with or without 'strategy', here manifest by the consumption of a limited resource to probabilistically influence the outcome of each round.

Several interesting features of the FF combat system make it potentially noteworthy with respect to similar systems (Canbolat et al., 2012; Neller & Presser, 2004; Smith, 2007). The allocation of resource is dynamic and depends on system state (Blanning, 1981; Hopp, 1987; Posner & Zuckerman, 1990). The use of resource can both increase the probability of a positive outcome for the player, or a negative outcome for the opponent. Use of this resource does not guarantee a positive outcome: its use is a gamble (Dubins & Savage, 1965; Maitra & Sudderth, 2012) that may negatively affect the player. The probability of this negative effect increases as more resource is used, providing an important consideration in the decision of whether to invoke this policy in the face of 'diminishing returns' (Deckro & Hebert, 2003).

This analysis reveals several structural properties of the system that are not specific to the FF context of this study. The case of a limited resource being gambled against to both amplify successes and mitigate failures (with the probability of a positive outcome diminishing as resource is used) bears some resemblance to, for example, a conceptual picture of espionage and counterespionage (Solan & Yariv, 2004). For example, resource could be allocated to covert operations with some probability of amplifying thefts and mitigating losses of information, with increased probability of negative outcomes due to discovery as these covert activites are employed more. Our analysis then reveals the strategies to be employed in different scenarios of information security (SKILL) and robustness to information loss (STAMINA). Another picture is one of cheating in games. Here, SKILL reflects the propensity to win a given round of the game without cheating, STAMINA reflects the number of losses before overall defeat, and LUCK a degree of covertness allows cheating without detection, but is diminished with repeated use.

In both of these scenarios, the coarse-grained optimal policy decision in Fig. 4 is informative. The decision whether or not to intervene (employ espionage, cheat, or so on) is based first on the probability of being detected – if high, interventions should be restricted to emergency situations where defeat is imminent. If low, the decision is next informed by the probability of winning individual rounds. If this is high, resource can be stored to hedge against future emergency situations. If not, the specific state of play informs how best to intervene – to amplify successes, mitigate defeats, or both.

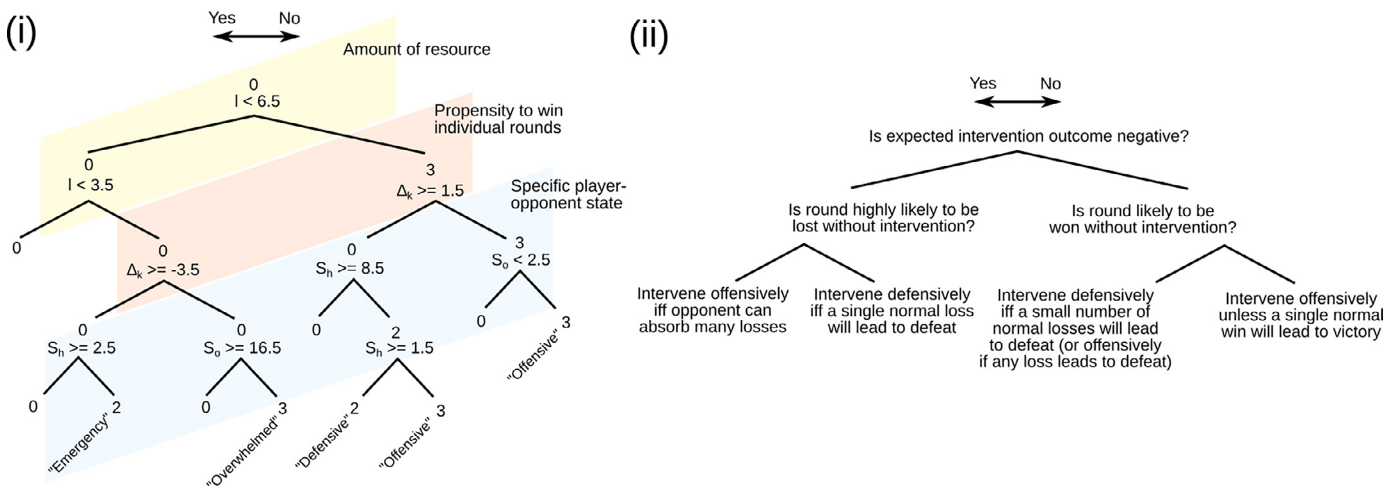

Fig. 4. Decision tree approximating optimal strategy choice. (i) Classification tree outputting a numerical code for an intervention strategy based on the current state of the system $\{\Delta_k, l, s_h, s_o\}$. 0 – do not use LUCK regardless of the outcome of this round; 2 – use LUCK if this round is lost (defensively); 3 – use LUCK if this round is won (offensively). For each inequality, go left for a positive response and right for a negative response. Top-level questions query amount of intervention resource (l), followed by individual victory propensity (Δ_k) and then specific state details. (ii) The more general phrasing of the decision strategy in (i) and in an alternative parameterisation of the game (see text), illustrating the coarse-grained general principles for approximating optimal policy.

In the absence of strategy, we find a closed-form expression for victory probability that takes an intuitive form. When strategy is included, dramatic increases in victory probability are found. The strong advantages provided by successful use of resource towards the 'endgame', where a successful gamble will produce instant victory or avoid instant defeat, shapes the structure of the optimal policy landscape. When little resource is available, complex structures emerge in the optimal landscape that depend on the tradeoff between using resource in the current state or 'saving' it in case of a more beneficial state later ('risking points is not the same as risking the probability of winning' Neller & Presser (2004)). When default victory is unlikely, using resource to reinforce rare success probabilities is a favoured strategy; when default victory is likely, using resource to mitigate rare loss probabilities is favoured. The specific optimal policy in a given state is solved and can be reasonably approximated by more heuristic strategies (Smith, 2007).

A route for expansion involves optimising victory probability while preserving some statistics, for example enforcing that $s_h > s^*$ or $l > l^*$ at victory, so that some resource is retained for subsequent tasks (for example, the rest of the adventure after this combat). Such constraints could readily be incorporated through an initial reallocating v_p over system states in the dynamic programming approach (Supplementary Fig. S1(i)), or by expanding the definition of the score being optimised to include some measure of desired retention in addition to v_p.

In an era of artificial intelligence approaches providing effective but essentially uninterpretable strategies for complex games (Campbell, Hoane, & Hsu, 2002; Gibney, 2016; Lee et al., 2016), more targetted analyses still have the potential to inform more deeply about the mechanisms involved in these strategies. Further, mechanistic understanding makes successful strategies readily available and simple to implement in the absence of computational resource. We believe that this analysis has demonstrated both some FF-specific and more general principles to the interesting case where a stochastic game can be influenced by gambling diminishing resource on interventions. In particular, we hope that the computation of optimal policy space and its subsequent coarse-grained representation may highlight general decision approaches for such systems. We hope that this increased interpretability and accessibility both contribute to the demonstration of the general power of these approaches, and also help improve the experience of some of the millions of FF players worldwide.

Acknowledgments

The author is grateful to Steve Jackson, Ian Livingstone, and the many other FF authors for creating this system and the associated worlds. The author also thanks Daniel Gibbs and Ellen Røyrvik for productive discussions.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejor.2022.01.039.

References

Baye, M. R., & Hoppe, H. C. (2003). The strategic equivalence of rent-seeking, innovation, and patent-race games. *Games and Economic Behavior, 44*(2), 217–226.
Bellman, R. (1957). A Markovian decision process. *Journal of Mathematics and Mechanics*, 679–684.
Bellman, R. (1965). On the application of dynamic programing to the determination of optimal play in chess and checkers. *Proceedings of the National Academy of Sciences of the United States of America, 53*(2), 244.
Blanning, R. W. (1981). Variable-base budgeting for R&D. *Management Science, 27*(5), 547–558.
Campbell, M., Hoane, A. J., Jr, & Hsu, F.-h. (2002). Deep blue. *Artificial Intelligence, 134*(1–2), 57–83.
Campbell, P. J. (2002). Farmer Klaus and the Mouse. *The UMAP Journal, 23*(2), 121–134.
Canbolat, P. G., Golany, B., Mund, I., & Rothblum, U. G. (2012). A stochastic competitive R&D race where "winner takes all". *Operations Research, 60*(3), 700–715.
Clark, C. W., & Harvell, C. D. (1992). Inducible defenses and the allocation of resources: A minimal model. *The American naturalist, 139*(3), 521–539.
Clark, D. J., & Riis, C. (1998). Contest success functions: An extension. *Economic Theory, 11*(1), 201–204.
Clarke, S. R., & Norman, J. M. (2003). Dynamic programming in cricket: Choosing a night watchman. *Journal of the Operational Research Society, 54*(8), 838–845.
Costikyan, G. (2007). Games, storytelling, and breaking the string. *Second Person: Roleplaying and Story in Playable Media*, 5–14.
Deckro, R. F., & Hebert, J. E. (2003). Modeling diminishing returns in project resource planning. *Computers & industrial engineering, 44*(1), 19–33.
Deutsch, Y. (2021). A polynomial-time method to compute all Nash equilibria solutions of a general two-person inspection game. *European Journal of Operational Research, 288*(3), 1036–1052.
Deutsch, Y., Golany, B., & Rothblum, U. G. (2011). Determining all Nash equilibria in a (bi-linear) inspection game. *European Journal of Operational Research, 215*(2), 422–430.

ARTICLE IN PRESS

Dresher, M. (1962). A sampling inspection problem in arms control agreements: A game-theoretic analysis. *Technical Report*. RAND CORP SANTA MONICA CA.
Dubins, L. E., & Savage, L. J. (1965). *Inequalities for stochastic processes: How to gamble if you must*. Dover Publications.
Gascoigne, M. (1985). *Out of the Pit*. Puffin Books.
Gerchak, Y., & Parlar, M. (1999). Allocating resources to research and development projects in a competitive environment. *IIE Transactions, 31*(9), 827–834.
Gibney, E. (2016). Google AI algorithm masters ancient game of Go. *Nature News, 529*(7587), 445.
Golany, B., Goldberg, N., & Rothblum, U. G. (2015). Allocating multiple defensive resources in a zero-sum game setting. *Annals of Operations Research, 225*(1), 91–109.
Golany, B., Kaplan, E. H., Marmur, A., & Rothblum, U. G. (2009). Nature plays with dice–terrorists do not: Allocating resources to counter strategic versus probabilistic risks. *European Journal of Operational Research, 192*(1), 198–208.
Green, J. (2014). *You are the Hero*. Snowbooks Ltd.
Heidenberger, K., & Stummer, C. (1999). Research and development project selection and resource allocation: A review of quantitative modelling approaches. *International Journal of Management Reviews, 1*(2), 197–224.
Hopp, W. J. (1987). A sequential model of R&D investment over an unbounded time horizon. *Management Science, 33*(4), 500–508.
Jackson, S., & Livingstone, I. (2002). *The warlock of firetop mountain*. Wizard Books.
Johnston, I. G. (2016). Endless love: On the termination of a playground number game. *Recreational Mathematics Magazine, 3*(5), 61–78.
Kallenberg, L. (2003). Finite state and action MDPs. In *Handbook of Markov decision processes* (pp. 21–87). Springer.
Kirk, D. E. (2012). *Optimal control theory: An introduction*. Courier Corporation.
Kostreva, M. M., & Hartman, R. (2003). Multiple objective solutions for Tetris. *Journal of Recreational Mathematics, 32*(3), 214.
Kuykendall, C., & Mackenzie, D. (1999). Analyzing solitaire. *Science, 283*(5403). 794–794
Lee, C.-S., Wang, M.-H., Yen, S.-J., Wei, T.-H., Wu, I.-C., Chou, P.-C., … Yan, T.-H. (2016). Human vs. computer go: Review and prospect [discussion forum]. *IEEE Computational Intelligence Magazine, 11*(3), 67–72.
Lee, J. D. (2012). The comparison of strategies used in the game of RISK via Markovian analysis and Monte-Carlo simulation. *Technical Report AFIT/IOA/ENS/12-02*. Air Force Inst of Tech Wright-Patterson AFB OH Graduate School of Engineering and Management.
Little, J. D. C., Murty, K. G., Sweeney, D. W., & Karel, C. (1963). An algorithm for the traveling salesman problem. *Operations Research, 11*(6), 972–989.
Maitra, A. P., & Sudderth, W. D. (2012). *Discrete gambling and stochastic games*: vol. 32. Springer Science & Business Media.
Neller, T. W., & Presser, C. G. M. (2004). Optimal play of the dice game pig. *The UMAP Journal, 25*(1).
OEIS (2019). The on-line encyclopedia of integer sequences, Published electronically at https://oeis.org.
Percy, D. F. (2015). Strategy selection and outcome prediction in sport using dynamic learning for stochastic processes. *Journal of the Operational Research Society, 66*(11), 1840–1849.
Perea, F., & Puerto, J. (2007). Dynamic programming analysis of the TV game "who wants to be a millionaire?". *European Journal of Operational Research, 183*(2), 805–811.
Posner, M., & Zuckerman, D. (1990). Optimal R&D programs in a random environment. *Journal of Applied Probability, 27*(2), 343–350.
Rabb, A. (1988). *A probabilistic analysis of the game of solitaire*. Harvard University Undergraduate honors thesis.
Smith, D. K. (2007). Dynamic programming and board games: A survey. *European Journal of Operational Research, 176*(3), 1299–1318.
Solan, E., & Yariv, L. (2004). Games with espionage. *Games and Economic Behavior, 47*(1), 172–199.
Team, R. C. (2018). R: A language and environment for statistical computing. R Foundation for Statistical ComputingVienna, Austria. https://www.R-project.org/.
Thomas, L. C. (2003). The best banking strategy when playing the weakest link. *Journal of the Operational Research Society, 54*(7), 747–750.
Trick, M. A. (2001). Building a better game through dynamic programming: A flip analysis. *INFORMS Transactions on Education, 2*(1), 50–58.
Valenzuela, M., Szidarovszky, F., & Rozenblit, J. (2015). A multiresolution approach for optimal defense against random attacks. *International Journal of Information Security, 14*(1), 61–72.