Markov chain analysis of overkill and natural regeneration in OSRS combat

Nukelawe

July 15, 2020

Contents

1 Introduction

Most activities in Old School RuneScape [1] are repetitive and feature randomness, making it important to optimize the related expected value quantities. In the case of combat, the subject of optimization is typically either the kill rate or the damage rate (DPS). A method for calculating these quantities is desirable. The main difficulty in these calculations is caused by overkill, which occurs when the enemy does not have enough hitpoints to receive all the damage that was rolled by the attaker. Extra complexity is added by natural regeneration, which affects almost every fight to some, albeit small extent. Earlier attempts at calculating the kill and damage rates have all made the assumption of no natural regeneration. An approximation for the damage rate was given by Nukelawe [7] using simplifying assumptions about the long term enemy hitpoint distribution. A more accurate model was presented by Palfore [8] and an attempt at an exact solution for the regenerationless case was made by Corpslayer [5] who also derived an asymptotic approximation [4].

In this study we will drop the assumption of no regeneration and treat Oldschool Runescape fights as absorbing Markov chains. This will allow us to obtain an exact solution for the expected number of hits to kill and the expected number of hitpoints regenerated. From these quantities, we will be able to calculate the kill and damage rates as long time averages. In search of less computationally intensive models and a better understanding of the solution we will additionally study in detail the case of no regeneration as well as some other approximations.

2 Fight mechanics

2.1 Damage

Consider a scenario in which a single attacker is fighting an unlimited supply of identical passive enemies one at a time. We denote the **maximum hitpoints** of an enemy by h. Once the remaining hitpoints of an enemy reaches 0 the enemy dies and is immediately replaced by a new one with all of its hitpoints remaining. During a fight, the enemy is hit periodically, once every **attack period** (denoted by T_A) and the damage dealt is calulated as follows.

1. The game determines if the hit will succeed by some random process, which depends on various parameters such as the defensive bonuses of the enemy, offensive bonuses of the attacker and appropriate combat-related stats. We ignore the details of this process and simply call the probability of success the **accuracy** (denoted by a). If the accuracy check fails, the attack is considered a miss and 0 damage is dealt.

2. If the accuracy check is successful, damage roll $M \sim U[0, m]$, a uniform random integer between 0 and m, is chosen, where m is the **maximum hit**. Again, we do not concern ourselves with the specifics of how the max hit is determined.

3. If $M > H$, where H is the remaining hitpoints of the enemy, the damage is capped to H and the final damage dealt is $\min(M, H)$. This is the overkill effect.

Notice that it is possible for the hit to deal no damage even if the accuracy check succeeds because the damage roll M could be 0. For the specifics of max hit and accuracy calculation see for example the OSRSwiki articles [2] on maximum hit and the OSRS combat overview [8] by Palfore.

2.2 Regeneration

Natural regeneration is a process which periodically attempts to increment the remaining hitpoints by one. A regeneration attempt will fail if the hitpoints are already full. The period T_R of the healing cycle is called the **regeneration period** and can vary between enemies giving rise to different regeneration rates. At the beginning of a fight the state of the healing cycle can be assumed to be unknown and thus, treated as a uniformly distributed random variable $t_0 \sim U[0, T_R - 1]$. This way only the timing of the first regeneration attempt is random and the rest are perfectly periodic.

If measured in ticks [3], the periods T_A and T_R as well as all other quantities that describe time can be assumed integers. We also assume $T_R > T_A$ which ensures that there will be at most one regeneration attempt between any two hits. This assumption should hold for nearly all cases in practice as the the typical regeneration period is 100 ticks (60 seconds) and even the slowest weapons have attack periods of just 7 ticks (4.2 seconds). Another assumption is that the damage dealt by a hit is calculated *after* regeneration if a regeneration attempt occurs on the same tick as the hit. Doing so avoids the edge case in which the enemy dies and then immediately regenerates back to life.

3 Markov model

We define the *state space* as the set of all possible values the remaining hitpoints of the enemy can have during a fight. For a fight against an enemy with h maximum hitpoints this is $\{0, \ldots, h\}$. As each hit corresponds to a transition from one state to another we can think of the fight as a random walk in the state space that terminates in state 0. As an example consider a 4-hitpoint enemy that the attacker hits 3 times: first 1, then 0 and finally 3 damage, killing the enemy. This fight would be described by the sequence $4, 3, 3, 0$ of visited states and corresponds to the highlighted path in the state space diagram in Figure 3.1.

The probability of transitioning from state i to state j is called the *transition probability* and defined as

$$p_{i,j} = \mathrm{P}(H_k{=}j \mid H_{k-1}{=}i) \tag{3.1}$$

where H_k is the number of hitpoints the enemy has remaining after k hits. The $h+1$ wide square matrix $\mathbf{T}$ that has the transition probabilities as its elements is called the *transition matrix*.

$$\mathbf{T} = \begin{pmatrix} p_{0,0} & p_{0,1} & \cdots & p_{0,h} \\ p_{1,0} & p_{1,1} & \cdots & p_{1,h} \\ \vdots & \vdots & \ddots & \vdots \\ p_{h,0} & p_{h,1} & \cdots & p_{h,h} \end{pmatrix} \tag{3.2}$$

The n-step transition probability is the probability of transitioning from state i to j in n steps, and we denote it by

$$p_{i,j}^{(n)} = \mathrm{P}(H_k{=}j \mid H_{k-n}{=}i). \tag{3.3}$$

It can be read off as the element i,j of the transition matrix power $\mathbf{T}^n$, that is $p_{i,j}^{(n)} = [\mathbf{T}^n]_{i,j}$.
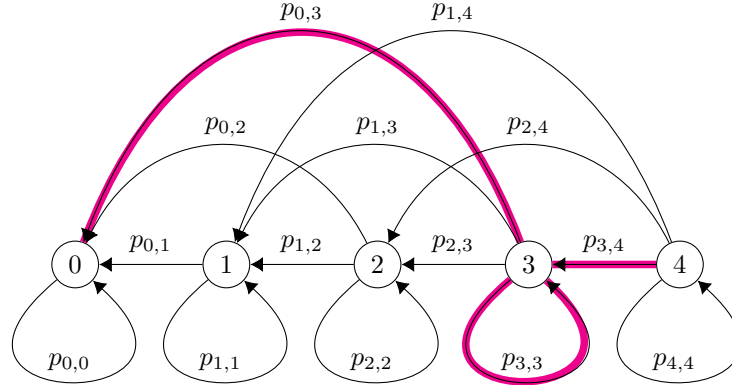

Figure 3.1: State space diagram for $m = 3$, $h = 4$ and no regeneration. The probability of moving from state i to j is given with the transition probability $p_{i,j}$. Only edges for which $p_{i,j} > 0$ are shown. The highlighted path corresponds to an example fight.

3.1 Transition probabilities

Let the accuracy-corrected damage roll X be the damage that would be dealt without considering overkill, or equivalently, on an enemy that had enough hitpoints to receive it. Using the fight mechanics presented in Chapter 2.1 its probability distribution can be worked out to be

$$\mathrm{P}(X{=}n) = \begin{cases} 1 - \frac{am}{m+1}, & \text{if } n = 0 \\ \frac{a}{m+1}, & \text{if } 1 \leq n \leq m \\ 0, & \text{if } n < 0 \text{ or } m < n \end{cases} \tag{3.4}$$

4

where a is the accuracy and m the maximum hit. By also considering overkill we obtain the regenerationless transition probabilities:

$$\pi_{i,j} = \begin{cases} P(X{=}i{-}j), & \text{if } j > 0 \\ P(X{\geq}i{-}j), & \text{if } j = 0 \end{cases}.$$ (3.5)

When regeneration is considered there are two ways to transition from state i to j. To lower the enemy hitpoints by n one can either heal 0 and inflict n damage, or heal 1 and inflict $n + 1$ damage. Which one happens depends on whether there is a successful regeneration attempt between the hits. Let $\mathcal{R}_k$ be the event that a regeneration attempt occurs between the $(k - 1)$th and the kth hits, and denote its complementary event by $\mathcal{R}_k^c$. Remembering that a regeneration attempt will only succeed if the enemy hitpoints are not full or zero we can write the transition probabilities as

$$P(H_k{=}j \mid H_{k-1}{=}i) = \begin{cases} P(\mathcal{R}_k)\pi_{i+1,j} + P(\mathcal{R}_k^c)\pi_{i,j}, & \text{if } 0 < i < h \\ \pi_{i,j}, & \text{if } i = h \text{ or } i = 0 \end{cases}.$$ (3.6)

To compute the probability $P(\mathcal{R}_k)$ consider the time until the next regeneration attempt immediately after the kth hit and denote it by t_k. Then t_0 is the time until the first regeneration attempt at the beginning of the fight, T_A ticks before the first hit. Since the initial state of the regeneration cycle is assumed to be unknown, t_0 is treated as a uniformly distributed random integer $t_0 \sim U[0, T_R - 1]$. Recall that t_0 can only be assumed an integer if T_R is given in ticks (0.6 seconds). The kth hit occurs after kT_A ticks from the beginning of the fight. Therefore,

$$t_k = (t_0 - kT_A) \bmod T_R$$ (3.7)

which is clearly also uniformly distributed $t_k \sim U[0, T_R - 1]$. Now the probability that there is a regeneration attempt between the $(k - 1)$th and the kth hit is

$$\rho = P(\mathcal{R}_k) = P(t_{k-1} < T_A) = \frac{T_A}{T_R}.$$ (3.8)

This number is called the *regeneration attempt frequency* and can also be interpreted as the probability of regeneration in a model in which regeneration has a chance of occurring between any two hits. Remarkably, it also does not depend on the hit number k at all, making the Markov process time-homogeneous.

In the end, the transition probabilities can be written as

$$p_{i,j} = \begin{cases} \rho\pi_{i+1,j} + (1 - \rho)\pi_{i,j}, & \text{if } 0 < i < h \\ \pi_{i,j}, & \text{if } i = h \text{ or } i = 0 \end{cases}$$ (3.9)

and the corresponding transition matrix has the form

$$\mathbf{T} = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{r} & \mathbf{Q} \end{pmatrix}$$ (3.10)

where $\mathbf{0} = (0, \ldots, 0)$ and $\mathbf{r}$ are a row vector and a column vector of length h, respectively, and $\mathbf{Q}$ is an h by h square matrix. Such a transition matrix is characteristic to an absorbing Markov chain. Its nth power, the elements of which are the n-step transition probabilities, is also easy to compute:

$$\mathbf{T}^n = \begin{pmatrix} 1 & \mathbf{0} \\ (\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \cdots + \mathbf{Q}^n)\mathbf{r} & \mathbf{Q}^n \end{pmatrix}$$ (3.11)

By carefully inserting different combinations of states into equation 3.9, the matrix $\mathbf{Q}$ can be observed to

have the following *almost* triangular form.

$$
\mathbf{Q} = \begin{pmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,h} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,h} \\ \vdots & \vdots & \ddots & \vdots \\ p_{h,1} & p_{h,2} & \cdots & p_{h,h} \end{pmatrix} = \begin{pmatrix} e_1 & e_2 \\ e_3 & e_1 & e_2 \\ \vdots & \vdots & \vdots & \ddots \\ e_3 & e_3 & e_3 & \cdots & e_2 \\ e_3 & e_3 & e_3 & \cdots & e_1 & e_2 \\ e_4 & e_3 & e_3 & \cdots & e_3 & e_1 & e_2 \\ & e_4 & e_3 & \cdots & e_3 & e_3 & e_1 & e_2 \\ & & e_4 & \cdots & e_3 & e_3 & e_3 & e_1 & e_2 \\ & & & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \\ & & & & e_4 & e_3 & e_3 & e_3 & e_3 & \cdots & e_2 \\ & & & & & e_4 & e_3 & e_3 & e_3 & \cdots & e_1 & e_2 \\ & & & & & & e_3 & e_3 & e_3 & \cdots & e_3 & e_5 \end{pmatrix} \tag{3.12}
$$

$$\underbrace{}_{m}$$

where

$$
e_1 = \frac{\rho a}{m+1} + (1-\rho)\left(1 - \frac{ma}{m+1}\right) \qquad\qquad e_2 = \rho\left(1 - \frac{ma}{m+1}\right)
$$

$$
e_3 = \frac{a}{m+1} \qquad\qquad e_4 = (1-\rho)\frac{a}{m+1}
$$

$$
e_5 = 1 - \frac{ma}{m+1}.
$$

3.2 Expected values

3.2.1 Number of visits

Let $N_{i,j}$ be the number of times the enemy has j hitpoints during a fight that starts from i hitpoints. The initial state is also counted, so for a fight that starts from full h hitpoints and never regenerates back to h, $N_{h,h}$ is the number of zeros hit at the beginning of the fight. We define the indicator function I_j^k such that, $I_j^k = 1$ if the enemy has j hitpoints after the kth hit and 0 otherwise. Now

$$
N_{i,j} = \sum_{k \geq 0} I_j^k \tag{3.13}
$$

and the expected number of visits to state j is

$$
\langle N_{i,j} \rangle = \left\langle \sum_{k \geq 0} I_j^k \right\rangle = \sum_{k \geq 0} \langle I_j^k \rangle = \langle I_j^0 \rangle + \sum_{k \geq 1} \langle I_j^k \rangle = \delta_{i,j} + \sum_{k \geq 1} p_{i,j}^{(k)} \tag{3.14}
$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 if $i \neq j$. Recall that $p_{i,j}^{(k)}$ is the element i,j of the matrix $\mathbf{T}^k$, specifically its transient submatrix $\mathbf{Q}^k$ since we are only considering $i, j > 0$. This allows us to write $\langle N_{i,j} \rangle$ as the element i, j of $\mathbf{I} + \mathbf{Q} + \mathbf{Q}^2 + \cdots$, where $\mathbf{I}$ is the identity matrix of the appropriate shape. Since the enemy will eventually die if hit infinitelly many times, $p_{i,j}^{(k)} \to 0 \iff \mathbf{Q}^k \to \mathbf{0}$ as $k \to \infty$. Therefore the infinite matrix series can, analogously to the regular geometric series, be expressed as

$$
\mathbf{I} + \sum_{k \geq 1} \mathbf{Q}^k = (\mathbf{I} - \mathbf{Q})^{-1}
$$

$$
\implies \langle N_{i,j} \rangle = \left[(\mathbf{I} - \mathbf{Q})^{-1} \right]_{i,j}. \tag{3.15}
$$

3.2.2 Number of hits to kill

Let a fight's length L_i be the number of hits to kill an enemy that has i hitpoints remaining. Then

$$
L_i = \sum_{j=1}^{h} N_{i,j} \implies \langle L_i \rangle = \sum_{j=1}^{h} \langle N_{i,j} \rangle. \tag{3.16}
$$

6

By defining the h-dimensional column vectors $\mathbf{L} = \left(\langle L_1 \rangle, \langle L_2 \rangle, \ldots, \langle L_h \rangle \right)^T$ and $\mathbf{1} = (1, 1, \ldots, 1)^T$, and using equation 3.15 this can be represented as a matrix equation

$$(\mathbf{I} - \mathbf{Q})\mathbf{L} = \mathbf{1} \tag{3.17}$$

which is equivalent to a linear system of h equations. The important special case $h = 1$ is particularly simple as it is unaffected by regeneration and the system of equations (3.17) reduces to a single equation:

$$(1 - p_{1,1}) \langle L_1 \rangle = 1$$
$$\implies \langle L_1 \rangle = \frac{1}{1 - p_{1,1}} = \frac{m+1}{am} \quad \text{if } h = 1. \tag{3.18}$$

3.2.3   Number of hitpoints regenerated

The number of hitpoints R_i regenerated during a fight against an enemy with i hitpoints remaining can be obtained by counting the number of successful regeneration attempts. An attempt is successful if the enemy has neither full nor 0 hitpoints at the time the attempt occurs. Let $\mathcal{S}_k$ denote the event that there is a successful regeneration attempt between the $(k-1)$th and the kth hits. Then by linearity of expectation

$$\langle R_i \rangle = \sum_{k \geq 0} \mathrm{P}(\mathcal{S}_k \mid H_0 = i)$$
$$= \sum_{k \geq 0} \rho \sum_{j=1}^{h-1} p_{i,j}^{(k)}$$
$$= \rho \sum_{j=1}^{h-1} \left(\delta_{i,j} + \sum_{k \geq 1} p_{i,j}^{(k)} \right)$$
$$= \rho \sum_{j=1}^{h-1} \langle N_{i,j} \rangle \tag{3.19}$$

Just like we did for $\langle L_i \rangle$, by defining the h-dimensional column vectors $\mathbf{R} = \left(\langle R_1 \rangle, \langle R_2 \rangle, \ldots, \langle R_h \rangle \right)^T$ and $\mathbf{1}_h = (1, 1, \ldots, 1, 0)^T$, equation 3.19 can be expressed in the matrix form

$$(\mathbf{I} - \mathbf{Q})\mathbf{R} = \rho \mathbf{1}_h. \tag{3.20}$$

Additionally, from equation 3.19 we can derive a handy approximation for $\langle R_h \rangle$. We begin by writing $\langle R_i \rangle$ as

$$\langle R_i \rangle = \rho \sum_{j=1}^{h-1} \langle N_{i,j} \rangle = \rho \left(\sum_{j=1}^{h} \langle N_{i,j} \rangle - \rho \langle N_{i,h} \rangle \right) = \rho \left(\langle L_i \rangle - \langle N_{i,h} \rangle \right)$$

Assume that the fight starts from full h hitpoints and the damage rate is high enough compared to the regeneration rate, that once left, the state h is never revisited. Then $\langle N_{h,h} \rangle$ can be approximated as the expected number of hits before the state h is left for the first time. This is equal to $\langle L_h \rangle$ in the case when $h = 1$, which was calculated to equal $\frac{m+1}{am}$. Therefore,

$$\boxed{\langle R_h \rangle \approx \rho \left(\langle L_h \rangle - \frac{m+1}{am} \right)} \tag{3.21}$$

3.3   Kill and damage rates

Typically the expected values discussed in Chapter 3.2 themselves are not the quantities of interest. Instead, the aim is often to optimize a rate quantity such as damage per second (DPS) or kills per hour. Such quantities can be easily expressed in terms of the expected values $\langle L \rangle$ and $\langle R \rangle$. Here we have dropped the index i assuming initial state is h, i.e. $\langle L \rangle \equiv \langle L_h \rangle$. We will be using the same notation also in the later chapters.

Consider a sequence of n fights against identical enemies. If the length of the ith fight in this sequence is denoted by L_i then the time taken by all the fights together is $T_A(L_1 + \cdots + L_n)$. Likewise, if the hitpoints regenerated during the ith fight is denoted by R_i then the total damage dealt is $nh + R_1 + \cdots + R_n$. The *kill rate* and *damage rate* are defined as

$$v_k = \lim_{n \to \infty} \frac{n}{T_A(L_1 + \cdots + L_n)} = \lim_{n \to \infty} \frac{1}{T_A \overline{L_n}} \tag{3.22}$$

$$v_d = \lim_{n \to \infty} \frac{nh + R_1 + \cdots + R_n}{T_A(L_1 + \cdots + L_n)} = \lim_{n \to \infty} \frac{h + \overline{R_n}}{T_A \overline{L_n}} \tag{3.23}$$

respectively, where $\overline{L_n} = \frac{1}{n}(L_1 + \cdots + L_n)$ is the average number of hits to kill an enemy and $\overline{R_n} = \frac{1}{n}(R_1 + \cdots + R_n)$ is the average hitpoints regenerated. Since both L_i and R_i are independent, identically distributed random variables, the law of large numbers implies that $\overline{L_n} \to \langle L \rangle$ and $\overline{R_n} \to \langle R \rangle$ as $n \to \infty$. Hence,

$$\boxed{v_k = \frac{1}{T_A \langle L \rangle} \quad \text{and} \quad v_d = \frac{h + \langle R \rangle}{T_A \langle L \rangle}.} \tag{3.24}$$

When T_A is in seconds, v_d is the DPS.

4 Approximations

In principle, the quantities $\langle L \rangle$ and $\langle R \rangle$ can now be solved from the matrix equations

$$(\mathbf{I} - \mathbf{Q})\mathbf{L} = \mathbf{1} \quad \text{and} \quad (\mathbf{I} - \mathbf{Q})\mathbf{R} = \rho \mathbf{1}_h$$

derived in Chapter 3. However, doing so requires solving systems of h linear equations, where h is the maximum hitpoints. This can be difficult or slow in practice, especially when h is large. In this chapter we will approximate the Markov process with certain simplifying assumptions.

4.1 No regeneration

Let us first assume that there is no regeneration ($\rho = \frac{T_A}{T_R} = 0$), which is a sensible thing to do since most of the time $T_A \ll T_R$. This reduces the transition matrix into a triangular one making an exact analytic solution using forward substitution possible. The effects of overkill still remain which makes such an approximation interesting.

We begin by restating the regenerationless transition probabilities

$$\pi_{i,j} = \begin{cases} P(X{=}i{-}j), & \text{if } j > 0 \\ P(X{\geq}i{-}j), & \text{if } j = 0 \end{cases} \tag{4.1}$$

where

$$P(X{=}n) = \begin{cases} 1 - \frac{am}{m+1}, & \text{if } n = 0 \\ \frac{a}{m+1}, & \text{if } 1 \leq n \leq m \\ 0, & \text{if } n > m \text{ or } n < 0. \end{cases} \tag{4.2}$$

Obviously, $\langle R_i \rangle = 0$ because there is no regeneration, so it suffices to find the expected length of a fight $\langle L_i \rangle$. Note that since there is now no way for the remaining hitpoints of the enemy to transition upwards, a fight against an enemy with i hitpoints remaining is equivalent to one against an enemy for which i is the maximum hitpoints. Therefore, we can without loss of generality assume that the initial state is h. The now triangular transition matrix allows solving for $\langle L_i \rangle$ recursively. To do this we write the matrix equation (3.17) explicitly.

$$\mathbf{1} = (\mathbf{I} - \mathbf{Q})\mathbf{L}$$

$$\implies 1 = \sum_{j=1}^{h} (\delta_{h,j} - \pi_{h,j}) \langle L_j \rangle$$

$$= (1 - \pi_{h,h}) \langle L_h \rangle - \sum_{j=1}^{h-1} \pi_{h,j} \langle L_j \rangle$$

$$= \big(1 - P(X{=}0)\big) \langle L_h \rangle - \sum_{j=1}^{h-1} P(X{=}h{-}j) \langle L_j \rangle$$

$$= \frac{am}{m+1} \langle L_h \rangle - \sum_{j=h-m}^{h-1} \frac{a}{m+1} \langle L_j \rangle$$

$$\implies \frac{am}{m+1} \langle L_h \rangle = 1 + \frac{a}{m+1} \sum_{j=h-m}^{h-1} \langle L_j \rangle$$

To simplify the summation limits, we have defined $\langle L_i \rangle = 0$ for $i \leq 0$. Finally, after reversing the order of summation and multiplying by $\frac{m+1}{am}$ the recurrence relation becomes

$$\langle L_h \rangle = \langle L_1 \rangle + \frac{1}{m} \sum_{j=1}^{m} \langle L_{h-j} \rangle \tag{4.3}$$

where $\langle L_1 \rangle = \frac{m+1}{am}$ as discussed in Chapter 3.2.2.

4.1.1 Analytic solution

Equation 4.3 is a linear recurrence of order m with constant coefficients. It can be solved using a generating function $g(z)$, which when expanded as power series around $z = 0$ has $\langle L_h \rangle$ as its coefficients:

$$g(z) = \sum_{h=1}^{\infty} \langle L_h \rangle \, z^h. \tag{4.4}$$

The generating function can be found by substituting the recurrence relation in place of $\langle L_h \rangle$ in equation 4.4.

$$g(z) = \sum_{h=1}^{\infty} \left(\frac{m+1}{am} + \frac{1}{m} \sum_{j=1}^{m} \langle L_{h-j} \rangle \right) z^h$$

$$= \frac{m+1}{am} \sum_{h=1}^{\infty} z^h + \frac{1}{m} \sum_{h=1}^{\infty} \sum_{j=1}^{m} \langle L_{h-j} \rangle \, z^h$$

$$= \frac{m+1}{am} \sum_{h=1}^{\infty} z^h + \frac{1}{m} \sum_{j=1}^{m} \sum_{h=1}^{\infty} \langle L_h \rangle \, z^{h+j}$$

$$= \frac{1}{m} \left(\frac{m+1}{a} \frac{z}{1-z} + \sum_{j=1}^{m} z^j g(z) \right)$$

$$= \frac{1}{m} \left(\frac{(m+1)z}{a(1-z)} + \frac{z - z^{m+1}}{1-z} g(z) \right)$$

$$\implies m(1-z)g(z) = \frac{(m+1)z}{a} + \left(z - z^{m+1} \right) g(z)$$

$$\implies \left(m - mz + z^{m+1} - z \right) g(z) = \frac{(m+1)z}{a}$$

$$\implies g(z) = \frac{(m+1)z}{a \left(m - (m+1)z + z^{m+1} \right)} \tag{4.5}$$

To read off the explicit formula for $\langle L_h \rangle$ the generating function must be expanded back into a power series. We begin by writing it in the form of a geometric sum.

$$g(z) = \frac{\frac{m+1}{am} z}{1 - \frac{z}{m}(m + 1 - z^m)}$$

$$= \frac{m+1}{am} z \sum_{n=0}^{\infty} \left(\frac{z}{m} \right)^n (m + 1 - z^m)^n$$

$$= \frac{m+1}{am} z \sum_{n=0}^{\infty} \left(\frac{z}{m} \right)^n \sum_{k=0}^{n} (m+1)^{n-k} (-1)^k z^{mk} \binom{n}{k}$$

$$= \frac{1}{a} \sum_{n=0}^{\infty} \sum_{k=0}^{n} \left(\frac{m+1}{m} \right)^{n+1} \left(\frac{-1}{m+1} \right)^k \binom{n}{k} z^{mk+n+1} \tag{4.6}$$

$\langle L_h \rangle$ is now the coefficient of the term z^h. Inspecting the exponent of z in equation 4.6 reveals that for the coefficient of interest the indices k and n must satisfy $h = mk + n + 1$. For each k there is exactly one valid n, namely $n = h - mk - 1$. Therefore, the coefficient of the hth order term in the series expansion of $g(z)$ is

$$\langle L_h \rangle = \frac{1}{a} \sum_{k \in \mathcal{I}} \left(\frac{m+1}{m} \right)^{h-mk} \left(\frac{-1}{m+1} \right)^k \binom{h - mk - 1}{k}$$

where $\mathcal{I}$ is the set of indices k such that

$$0 \le k \le n = h - mk - 1 \iff (m+1)k \le h - 1 \iff k \le \frac{h-1}{m+1}.$$

Since k must also be an integer the index set becomes $\mathcal{I} = \left\{ 0, 1, \ldots, \left\lfloor \frac{h-1}{m+1} \right\rfloor \right\}$ and the explicit solution to the recurrence can be written as

$$\langle L \rangle = \frac{1}{a} \sum_{k=0}^{\left\lfloor \frac{h-1}{m+1} \right\rfloor} \left(\frac{m+1}{m} \right)^{h-mk} \left(\frac{-1}{m+1} \right)^{k} \binom{h - mk - 1}{k} \tag{4.7}$$

where the index h in $\langle L_h \rangle$ has been omitted as the fight is assumed to start from full hitpoints. For small h equation 4.7 simplifies to the geometric sequence. In particular,

$$\langle L \rangle = \frac{1}{a} \left(\frac{m+1}{m} \right)^{h} \quad \text{if } h \leq m+1. \tag{4.8}$$

Note how the accuracy dependency can easily be factored out from equation 4.7. This gives a nice interpretation for $a \langle L \rangle$ as the number of *successful hits*, i.e. hits that pass the accuracy check.

4.1.2   Asymptotic approximation

Equation 4.7 can be annoying to work with, especially for large initial hitpoints h, so an approximation for the case $h \gg m$ is needed. Intuitively the length of a fight should have linear dependency on hitpoints as $h \to \infty$. This is because on average we can expect each hit to deal $\frac{am}{2}$ damage and therefore it should take approximately $\frac{2h}{am}$ hits to kill an enemy with h hitpoints. Due to overkill however, the last hit will tend to deal slightly less than $\frac{am}{2}$ damage, introducing a constant correction term to this linear approximation. We will now determine the magnitude of this constant by finding the asymptotic expansion for $\langle L_h \rangle$.

   To study the asymptotics we inspect the poles of the generating function which are simply the zeros of its denominator. Factorizing out the $(z-1)$ terms in the denominator allows writing the generating function as

$$g(z) = \frac{(m+1)z}{a\left(m - (m+1)z + z^{m+1}\right)} = \frac{(m+1)z}{a(z-1)^2 Q(z)} \quad \text{where} \quad Q(z) = \sum_{k=0}^{m-1} (m-k)z^{k}. \tag{4.9}$$

It is evident from equation 4.9 that the generating function has a second order pole at $z = 1$. The other poles q_k are the zeros of the polynomial $Q(z)$ and cannot be found exactly. However, since $Q(z)$ is a polynomial with positive coefficients, the Eneström-Kakeya theorem [6] can be used to obtain a lower bound

$$|q_k| \geq \max_{0 \leq k < m} \left(\frac{m-k}{m - (k+1)} \right) = \frac{m}{m-1} > 1. \tag{4.10}$$

Now, assume that for some k, q_k is not a simple root and thus has a multiplicity greater than 1. Then also the derivative of the denominator of $g(z)$ must have a zero at q_k, that is

$$(m+1)q_k^m - (m+1) = 0 \implies q_k^m = 1 \implies |q_k| = 1.$$

This contradicts the lower bound we just found (equation 4.10) implying that the poles q_k are all simple.

   Having discussed the properties of its poles we now express the generating function as a partial fraction

$$g(z) = \frac{\alpha}{(z-1)^2} + \frac{\beta}{z-1} + \sum_k \frac{\gamma_k}{q_k - z} \tag{4.11}$$

where γ_k are complex number constants and

$$\alpha = \lim_{z \to 1} (z-1)^2 g(z) = \lim_{z \to 1} \frac{(m+1)z}{aQ(z)} = \frac{m+1}{aQ(1)}, \tag{4.12}$$

$$\beta = \lim_{z \to 1} \frac{\mathrm{d}}{\mathrm{d}z}\left((z-1)^2 g(z)\right) = \lim_{z \to 1} \frac{m+1}{aQ(z)} \left(1 - \frac{zQ'(z)}{Q(z)} \right) = \alpha \left(1 - \frac{Q'(1)}{Q(1)} \right). \tag{4.13}$$

The values of the polynomial $Q(z)$ and its derivative $Q'(z)$ at $z = 1$ evaluate to $Q(1) = \frac{1}{2}m(m+1)$ and $Q'(1) = \frac{1}{6}m(m+1)(m-1)$. Substituting them back into equations 4.12 and 4.13 yields

$$\alpha = \frac{2}{am} \quad \beta = \frac{2(4-m)}{3am}. \tag{4.14}$$

Inserting α and β into equation 4.11 and expanding each term as a power series around $z = 0$ gives

$$g(z) = \frac{2}{am} \frac{1}{(1-z)^2} + \frac{2(m-4)}{3am} \frac{1}{1-z} + \sum_k \frac{\gamma_k}{q_k} \frac{1}{1 - \frac{z}{q_k}}$$

$$= \frac{2}{am} \sum_{n=0}^{\infty} (n+1)z^n + \frac{2(m-4)}{3am} \sum_{n=0}^{\infty} z^n + \sum_k \frac{\gamma_k}{q_k} \sum_{n=0}^{\infty} \frac{z^n}{q_k^n}$$

$$= \sum_{n=0}^{\infty} \left(\frac{2}{am}\left(n + \frac{m-1}{3}\right) + \sum_k \frac{\gamma_k}{q_k^{n+1}} \right) z^n$$

Remembering that $\langle L \rangle$ is the coefficient of the hth order term in the series expansion of the generating function we get

$$\langle L \rangle = \frac{2}{am}\left(h + \frac{m-1}{3} + \sum_k \frac{\gamma_k}{q_k^{h+1}} \right) \tag{4.15}$$

which is the asymptotic expansion we were looking for. By using the triangle inequality and equation 4.10 the last term can be bounded as follows.

$$\left| \sum_k \frac{\gamma_k}{q_k^{h+1}} \right| \leq \sum_k \frac{|\gamma_k|}{|q_k|^{h+1}} \leq \sum_k |\gamma_k|\left(\frac{m-1}{m}\right)^{h+1} = \left(\frac{m-1}{m}\right)^{h+1} \sum_k |\gamma_k| \tag{4.16}$$

Since $\sum_k |\gamma_k|$ is just a positive real number, we get $\sum_k \frac{\gamma_k}{q_k^{h+1}} = \mathcal{O}\left(\left(\frac{m-1}{m}\right)^h\right)$. Dropping the exponentially vanishing term yields the asymptotic approximation, which Corpslayer [4] also obtained using a different approach.

$$\boxed{\langle L \rangle \sim \frac{2}{am}\left(h + \frac{m-1}{3} \right)} \tag{4.17}$$

Figure 4.1 illustrates the quality of the asymptotic approximation for various values of h and m.
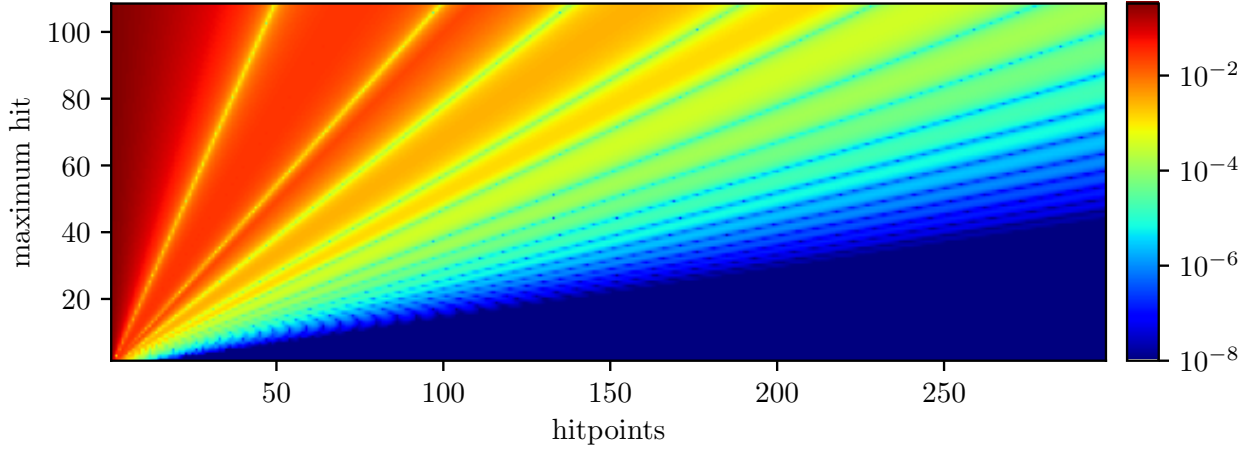

Figure 4.1: Magnitude of the relative error term $\frac{1}{\langle L \rangle} \sum_k \gamma_k q_k^{-h-1}$ for various values of maximum hit m and initial hitpoints h. As expected, the best accuracy is obtained when h is large compared to m. The complex roots q_k cause the asymptotic approximation to oscillate around the exact solution making the error term alternate in sign. The clearly visible lines of smaller error are where the sign flips.

4.2 Effective hitpoints approximation

Effects of regeneration can be added to the no regeneration approximation relatively easily by noticing that a fight against an enemy with h hitpoints is equivalent to one against an enemy with $h + R$ hitpoints and no regeneration, where R is the number of hitpoints regenerated during the fight. Since the expected damage of a hit that does not kill the enemy is $\frac{am}{2}$, the expected number of hits it takes to lower the enemy hitpoints by the extra R hitpoints is approximately $\frac{2R}{am}$. A naive attempt at approximating the effects of regeneration could be to assume that the same holds for the expectation:

$$\langle L \rangle \approx \langle L' \rangle + \frac{2}{am} \langle R \rangle \tag{4.18}$$

where $\langle L' \rangle$ is the length of the fight assuming no regeneration. Now, applying the approximation $\langle R \rangle \approx \rho\big(\langle L \rangle - \frac{m+1}{am} \big)$ obtained in Chapter 3.2 allows solving for $\langle L \rangle$ in terms of its regenerationless counterpart $\langle L' \rangle$.

$$\langle L \rangle \approx \langle L' \rangle + \frac{2\rho}{am} \Big(\langle L \rangle - \frac{m+1}{am} \Big)$$

$$\implies \frac{am - 2\rho}{am} \langle L \rangle \approx \langle L' \rangle - \frac{2\rho}{am} \frac{m+1}{am}$$

Multiplying both sides by $\frac{am}{am - 2\rho} = \frac{1}{1 - \frac{2\rho}{am}}$ gives

$$\boxed{\langle L \rangle \approx \frac{1}{1 - \frac{2\rho}{am}} \Big(\langle L' \rangle - \frac{2\rho(m+1)}{(am)^2} \Big)} \tag{4.19}$$

from which the number of hitpoints regenerated can be calculated as

$$\langle R \rangle \approx \rho\Big(\langle L \rangle - \frac{m+1}{am} \Big). \tag{4.20}$$

As a sanity check we note that if there is no regeneration and $\rho = 0$, then $\langle L \rangle = \langle L' \rangle$ and $\langle R \rangle = 0$ as they should. The quantity $\frac{2\rho}{am}$ that appears in equation 4.19 is called *relative regeneration rate* as it can be understood as the ratio between the rates at which the enemy hitpoints are incremented (ρ) and decremented ($\frac{am}{2}$). A relative regeneration rate of 1 means that, on average, the enemy hitpoints stays the same and the only way for death to ever occur is through random fluctuations. In general, we expect the model to only give good results when $\frac{2\rho}{am} \ll 1$. In contrast to the regenerationless case, the accuracy dependency of $\langle L \rangle$ can no longer be factored out, indicating that the number of hits and the number of *successful* hits are no longer trivially related when regeneration is taken into account.

4.2.1 Error analysis

Let us denote the relative error of quantity X by

$$\delta X = \frac{X - X'}{X}, \tag{4.21}$$

where the primed variable X' stands for the approximation of X. To study how well the effective hitpoints approximation holds in various regions of the parameter space, we will now inspect the relative errors of the kill rate v_k and the damage rate v_d. The true values of $\langle L \rangle$ and $\langle R \rangle$ needed for this comparison can be obtained by numerically solving the matrix equations 3.17 and 3.20.

Figure 4.2 shows how the relative errors of the rates v_k and v_d depend on m and h while keeping the other parameters a, T_A and T_R constant. First, we note that for both rates, the error is 0 for $h = 1$ because enemies with 1 hitpoint cannot regenerate. In case of the damage rate v_d, the error also vanishes when $m = 1$ because then there is no overkill and the expected damage per hit never changes no matter how long the fight lasts or how many hitpoints are regenerated. In the regions in which some error is present, the error is the largest near the origin where both m and h are small. This is particularly convenient since it is precisely in this region where the numerical solution to the exact Markov model is the fastest. There is currently no explanation for why the error seems to be higher on a blob around the main diagonal as well as on the line of slope $\frac{1}{2}$ intersecting the origin.
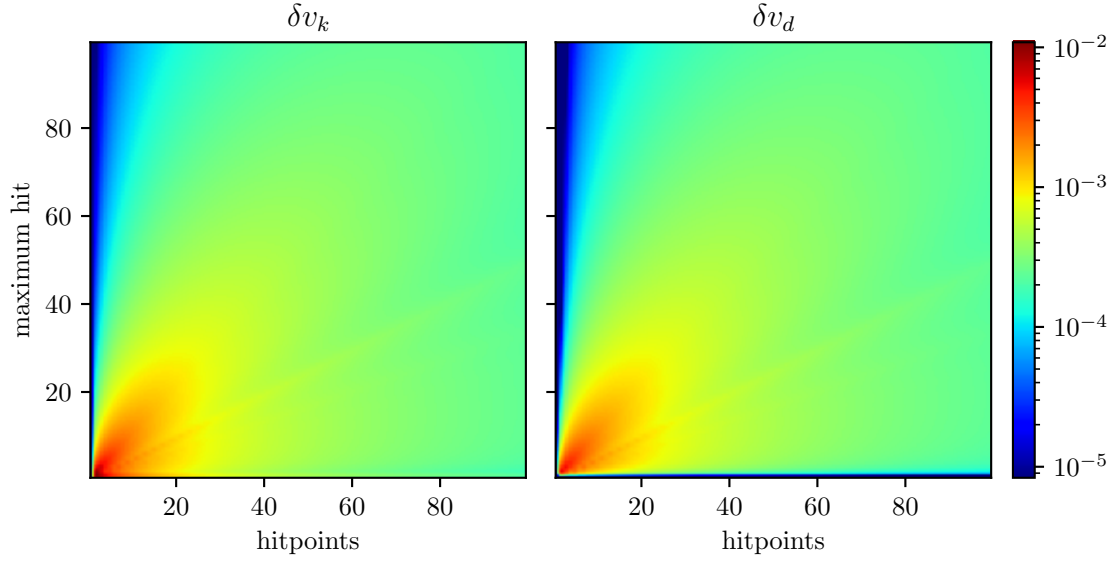
Figure 4.2: Relative errors of the kill rate v_k and the damage rate v_d in the effective hitpoint approximation for various values of maximum hit and enemy hitpoints. The other relevant parameters $T_A = 4$, $T_R = 100$ and $a = 0.5$ are held constant.
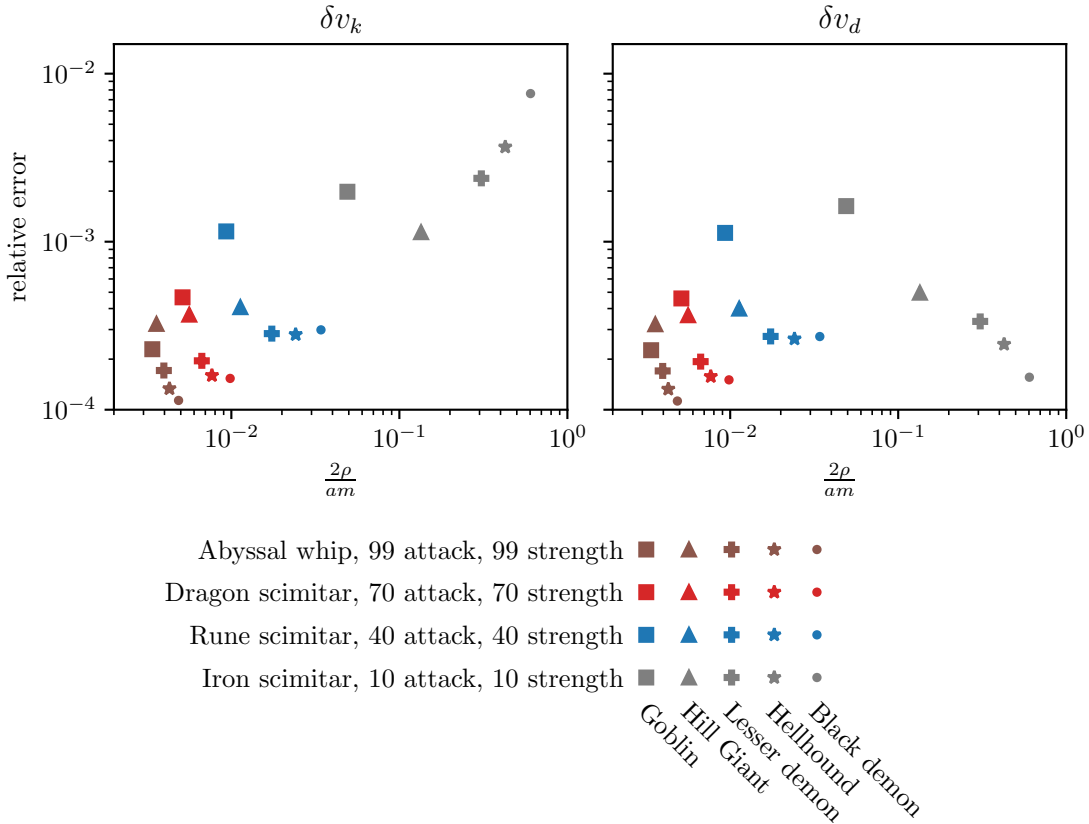

Figure 4.3: Relative errors of the kill rate v_k and damage rate v_d in the effective hitpoint approximation as functions of the relative regeneration rate $\frac{2\rho}{am}$. The regeneration period is assumed to be $T_R = 100$ in each fight. The parameters a, m, h and T_A for the fight scenarios were calculated using the OSRSmath [9] python-library.

Figure 4.3 shows the relative errors of the kill and damage rates against the relative regeneration rate for various enemy and attacker combinations. For simplicity, no boosts of any kind are considered, and the only gear used is a weapon. The errors are all below 1% (10^{-2}), most even under 0.1% (10^{-3}). They would in reality only get smaller as more offensive gear and other boosts are added. The reason why the relative error of the kill rate increases as the relative regeneration rate approaches 1 is that equation 4.19 has a singularity at the point $\frac{2\rho}{am} = 1$. However, most of the fight scenarios in Figure 4.3 have relative regeneration rates of less than 0.1 placing them far from the singularity. The only examples in which the relative regeneration rate is anywhere near 1 are the ridiculous ones where a player with level 10 attack and strength, equipped with an iron scimitar, is fighting a relatively strong enemy such as a hellhound. It is no surprise that for such fights regeneration is more significant.

4.3 On transition matrix triangularization

Finally, we will briefly discuss another approximation for slow regeneration without going into too much detail. The convenience of the no regeneration assumption was ultimately based on the fact that it made the transition matrix triangular. The same can be accomplished by another, weaker assumption than abandoning regeneration altogether. To ensure that the enemy hitpoints never increase, it is enough to assume that a regeneration attempt never coincides with a hit that inflicts 0 damage.

The transition probabilities in this case are

$$p_{i,j} = \begin{cases} \rho\pi_{i+1,j} + (1-\rho)\pi_{i,j}, & \text{if } 0 < j \le i < h \\ \pi_{i,j}, & \text{if } i = h \text{ or } i = 0 \text{ or } j > i \end{cases} \tag{4.22}$$

which yields the recurrence relation

$$\langle L_i \rangle = \frac{1}{m-\rho}\left(\frac{m+1}{a} + \sum_{j=1}^{m} \langle L_{i-j} \rangle + \rho \langle L_{i-m-1} \rangle \right) \tag{4.23}$$

Although solving equation 4.23 is significantly harder than without regeneration, it can still be done to a certain extent and at the very least it provides a computationally faster way to solve for the effects of regeneration. Just like in the regenerationless case we find that for small h the fight's length follows a geometric progression:

$$\langle L \rangle = \frac{m+1}{am}\left(\frac{m-\rho+1}{m-\rho} \right)^{h-1} \quad \text{if } h \le m+1. \tag{4.24}$$

The main disadvantage to this approximation is that the effective hitpoints argument is no longer straightforward to make. Combined with a significantly more difficult to solve recurrence relation, this approximation seems less appealing.

5 Conclusions

We described the damage and natural regeneration mechanics in their full generality using an absorbing Markov chain. The quantities $\langle L \rangle$ (number of hits to kill) and $\langle R \rangle$ (number of hitpoints regenerated) were then shown to be solutions to linear systems of as many equations as the enemy has maximum hitpoints. Because solving such linear systems is computationally intensive for enemies with high hitpoints, we derived a variety of approximations, mainly focusing on small regeneration rates.

Treating in detail the special case of no regeneration we were able to get a good understanding of the effects of overkill. We derived an exact analytic solution (equation 4.7) for the number of hits to kill the enemy and also found a convenient asymptotic approximation (equation 4.17) for when enemy hitpoints is large compared to the attacker maximum hit. By applying the so called effective hitpoints argument, the regenerationless case was then extended to also approximately account for regeneration. For practically all realistic combat scenarios, the relative errors of the resulting effective hitpoints approximation were observed to be well under 1%, mostly under 0.1%. A more rigorous application of the effective hitpoints argument is still worth considering in the future as it might produce even better approximations.

We also briefly discussed a less restrictive assumption that still succeeds in triangularizing the transition matrix. However, because the resulting recurrence relations were more cumbersome to handle and the effective hitpoints argument was difficult to apply, we did not pay much attention to this approximation. Future work might focus on this or other ways of triangularizing the transition matrix. Some thought could also be given on generalizing the asymptotic approximation (equation 4.17) to consider regeneration.

To conclude, we note that the treatment of the relative simple combat mechanics in Oldschool Runescape can be quite involved. There are also plenty of other combat mechanics that could be considered such as fights against multiple enemies at once, fights against enemies that attack back, poison damage and special attacks.

References

[1] https://oldschool.runescape.com/.

[2] https://oldschool.runescape.wiki/w/Maximum_hit.

[3] https://oldschool.runescape.wiki/w/RuneScape_clock.

[4] Corpslayer. https://imgur.com/a/6613Tlu, 2019.

[5] Corpslayer. Overkill dps formulas. https://www.reddit.com/r/2007scape/comments/bcq3mj/overkill_dps_formulas, 2019.

[6] S. Kakeya. On the limits of the roots of an algebraic equation with positive coefficients. *Tohoku Mathematical Journal, First Series*, 2:140–142, 1912.

[7] Nukelawe. Effects of overkill in damage output in oldschool runescape. https://imgur.com/aykEahg, 2016.

[8] Palfore. The mathematics and optimization of osrs combat. https://www.reddit.com/r/2007scape/comments/faz5et/the_mathematics_of_osrs_combat/, 2020.

[9] Palfore. Osrsmath. https://github.com/Palfore/OSRSmath, 2020.