Oracle Rating System - Mathematical Explanation

1. Bradley-Terry Model Foundation

1.1 Core Probability Formula

The Bradley-Terry model assigns each player a **strength parameter** $s_i > 0$. The probability that player i beats player j is:

$$P(i \text{ beats } j) = \frac{s_i}{s_i + s_j}$$

Key Properties:

- Symmetric: $P(i \text{ beats } j) + P(j \text{ beats } i) = 1$
- If $s_i = s_j$, then $P = 0.5$ (equal strength)
- If $s_i \gg s_j$, then $P \approx 1$ (player i dominates)
- If $s_i \ll s_j$, then $P \approx 0$ (player i loses)

1.2 Why Bradley-Terry?

Unlike Elo (which uses a logistic function), Bradley-Terry:

- Has a clear probabilistic interpretation
- Relates directly to win ratios
- Scales naturally to multiple comparisons
- Can be extended to team games and draws

---

2. Keeper as Reference Point

2.1 Setting the Benchmark

We fix **Keeper's strength** at:

$$s_{\text{Keeper}} = 1.0$$

This serves as our reference point (like how Celsius fixes water's freezing point at 0°C).

Your requirement: Keeper's rating = 500 (middle of 0-1000 scale)

2.2 Win Probability Against Keeper

For any player with strength s:

$$P(\text{beat Keeper}) = \frac{s}{s+1}$$

Examples:

- If $s = 0.001$: $P = 0.001/(1.001) \approx 0.1\%$ (nearly hopeless)
- If $s = 1.0$: $P = 1/2 = 50\%$ (equal to Keeper)
- If $s = 999$: $P = 999/1000 = 99.9\%$ (nearly unbeatable)

---

3. Rating Scale Transformation

3.1 Desired Mapping

We want the rating R to directly represent win probability:

$$R = 1000 \times P(\text{beat Keeper})$$

This gives us:

- Rating 0 → 0% win rate vs Keeper
- Rating 500 → 50% win rate vs Keeper
- Rating 1000 → 100% win rate vs Keeper

3.2 Converting Strength to Rating

Given strength s and $s_{\text{Keeper}} = 1$:

$$P = \frac{s}{s+1}$$

$$R = 1000 \times \frac{s}{s+1}$$

Implementation:

```c
double rating_bt_to_scale(double s, double keeper_s) {
    double p_win = s / (s + keeper_s);
    return 1000.0 * p_win;
}
```

3.3 Converting Rating to Strength (Inverse)

Given rating R, find strength s:

$$R = 1000 \times \frac{s}{s+1}$$

$$\frac{R}{1000} = \frac{s}{s+1}$$

Let $P = R/1000$:

$$P(s+1) = s$$

$$Ps + P = s$$

$$P = s - Ps = s(1-P)$$

$$s = \frac{P}{1-P}$$

Final formula:

$$s = \frac{R/1000}{1 - R/1000} = \frac{R}{1000 - R}$$

Examples:

- $R = 0$: $s = 0/(1000) = 0$
- $R = 500$: $s = 500/500 = 1$ ✓ (equals Keeper)
- $R = 750$: $s = 750/250 = 3$ (3× stronger than Keeper)
- $R = 900$: $s = 900/100 = 9$ (9× stronger than Keeper)

---

4. Updating Ratings After Matches

4.1 The Learning Problem

After observing match results, we need to update strength estimates. The Bradley-Terry model maximizes the **log-likelihood**:

$$\mathcal{L} = \sum_{(i,j)\in\text{matches}} w_{ij} \log\left(\frac{s_i}{s_i + s_j}\right)$$

where w_{ij} is the number of times i beat j.

4.2 Gradient Ascent Update

The gradient with respect to s_i is:

$$\frac{\partial \mathcal{L}}{\partial s_i} = \sum_j \left(\frac{w_{ij}}{s_i} - \frac{w_{ij} + w_{ji}}{s_i + s_j}\right)$$

This is complex for batch updates, so we use **incremental updates** after each match.

4.3 Incremental Update (Elo-style)

After player i plays n games against player j with score S_i (wins + 0.5×draws):

Actual score: $\bar{S}_i = S_i/n$ (ranges from 0 to 1)

Expected score: $E_i = P(i \text{ beats } j) = \frac{s_i}{s_i+s_j}$

Update rule: Use multiplicative update to keep strengths positive:

$$s_i^{\text{new}} = s_i \times \exp(k \times (\bar{S}_i - E_i))$$

where k is the learning rate (K-factor).

Why exponential?

- Ensures $s_i > 0$ always
- Small errors: $\exp(x) \approx 1 + x$ (approximately additive)
- Large errors: prevents negative strengths

4.4 Maintaining Keeper as Reference

After any update involving Keeper:

1. Renormalize: $s_{\text{Keeper}} = 1.0$
2. Scale other player proportionally if needed

This prevents **rating inflation/deflation** over time.

---

5. Win Probability Between Any Two Players

5.1 Direct Calculation

Given players i and j with ratings R_i and R_j:

Convert to strengths:

$$s_i = \frac{R_i}{1000 - R_i}, \quad s_j = \frac{R_j}{1000 - R_j}$$

Calculate probability:

$$P(i \text{ beats } j) = \frac{s_i}{s_i + s_j}$$

5.2 Example Calculation

Player A: Rating 600

Player B: Rating 450

$$s_A = \frac{600}{400} = 1.5$$

$$s_B = \frac{450}{550} = 0.818$$

$$P(A \text{ beats } B) = \frac{1.5}{1.5 + 0.818} = \frac{1.5}{2.318} = 0.647$$

Player A has a 64.7% chance of beating Player B.

5.3 Sanity Checks

Test 1: Player at 500 vs Keeper (500)

- $s_1 = 500/500 = 1$
- $s_K = 1$
- $P = 1/(1+1) = 0.5$ ✓

Test 2: Player at 750 vs Keeper (500)

- $s_1 = 750/250 = 3$
- $s_K = 1$
- $P = 3/(3+1) = 0.75$ ✓ (matches 75% expected)

---

6. Confidence Intervals

6.1 Uncertainty Estimation

With limited games, ratings are uncertain. We estimate using **Wilson score interval**:

For n games with win rate $\hat{p} = \text{wins}/n$:

$$\text{Standard Error} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

95% Confidence Interval (z = 1.96):

$$\text{CI} = 1.96 \times \text{SE} \times 1000$$

6.2 Convergence

As $n \to \infty$:

- SE $\to 0$
- Confidence interval narrows
- Rating becomes more reliable

Example:

- After 10 games: CI ≈ ±100 points
- After 50 games: CI ≈ ±50 points
- After 200 games: CI ≈ ±25 points

---

7. Mathematical Properties

7.1 Transitivity

Bradley-Terry is **not perfectly transitive**:

- If $P(A \text{ beats } B) = 0.7$ and $P(B \text{ beats } C) = 0.7$
- Then $P(A \text{ beats } C) \neq 0.49$ (it depends on relative strengths)

Strength-based transitivity:

$$s_A/s_B = r_1, \quad s_B/s_C = r_2 \implies s_A/s_C = r_1 \cdot r_2$$

7.2 Scale Invariance

Multiplying all strengths by constant $\alpha > 0$ doesn't change probabilities:

$$P = \frac{s_i}{s_i + s_j} = \frac{\alpha s_i}{\alpha s_i + \alpha s_j}$$

This is why we fix Keeper at 1.0 — to break scale ambiguity.

7.3 Rating Differences

Unlike Elo (where rating difference determines probability), Bradley-Terry uses **strength ratios**:

$$\frac{s_i}{s_j} = \frac{R_i(1000 - R_j)}{R_j(1000 - R_i)}$$

Rating difference of 100 points:

- 500 vs 400: s ratio = 1.25
- 700 vs 600: s ratio = 1.75
- 900 vs 800: s ratio = 4.5

The same rating gap means more at higher ratings!

---

8. Comparison with Elo

Feature	**Elo**	**Bradley-Terry (Our System)**
Win Probability	Logistic: $\frac{1}{1+10^{-\Delta/400}}$	Ratio: $\frac{s_i}{s_i+s_j}$
Scale	Arbitrary (1500 typical)	Fixed to win rate (0-1000)
Update	$R_{\text{new}} = R + K(S - E)$	$s_{\text{new}} = s \cdot e^{k(S-E)}$
Interpretation	Abstract	Direct probability
Transitivity	Better preserved	Strength-based

---

9. Practical Implementation Notes

9.1 Numerical Stability

Problem: Extreme ratings (0 or 1000) cause division issues.

Solution: Clamp ratings to [0.001, 999.999] before converting to strength.

9.2 K-Factor Tuning

- **High K** (32-64): Fast adaptation, more volatile
- **Low K** (8-16): Slow adaptation, more stable
- **Adaptive K**: Decrease with game count

9.3 Initial Rating

New players start at **500** (equal to Keeper) with high uncertainty. This is:

- Conservative (doesn't assume strength)
- Centers the distribution
- Converges quickly with data

---

10. Example Scenarios

Scenario A: New Human Player

Games vs Keeper:

1. Win 3/10 → Rating drops to ~350
2. Win 5/10 → Rating adjusts to ~480
3. Win 7/10 → Rating climbs to ~620

Scenario B: Three-Way Tournament

- Agent A: 650 (beats Keeper 65%)
- Agent B: 500 (equals Keeper)
- Agent C: 350 (beats Keeper 35%)

Predicted matchups:

- $P(A \text{ beats } B) = 1.86/2.86 = 65\%$
- $P(B \text{ beats } C) = 1.0/1.54 = 65\%$
- $P(A \text{ beats } C) = 1.86/2.40 = 78\%$

The system naturally handles multi-way comparisons!

Summary

The rating system uses:

1. **Bradley-Terry strengths** for probabilistic modeling
2. **Keeper (s=1)** as fixed reference point
3. **Linear scale** where rating = 1000 × P(beat Keeper)
4. **Gradient-based updates** for learning from matches
5. **Confidence intervals** for uncertainty quantification

This gives you a mathematically sound, interpretable rating system where every rating directly corresponds to expected performance against the Keeper benchmark.