Diploma Thesis

An Intelligent Artificial Player for the
Game of Risk

Michael Wolf
Agenda

• Risk
• Risk Framework
• Basic Evaluation Player (BEP)
• Enhanced Evaluation Player (EEP)
• Learning Player (LP)
• Human Opponents
• Conclusion
Risk
Risk – Overview

• Popular Strategy Board Game
• 2-6 Players
• Variable Move Sequence

• World / Continents / Territories
• Armies
• Conquer the World
Risk – Game Flow

1. Trading in Risk Cards
2. Placing new Armies
3. Attacking
4. Fortifying the Position
Risk – Complexity

• State-Space Complexity

• Game-Tree Complexity
  – Branching Factor ^ Game Turns

• In Risk: Infinite

→ Estimates Needed
  → State-Space Complexity: Cap on Army Number on Game Board
  → Game-Tree Complexity: Measure Test Games
## Risk – Complexity Comparison

<table>
<thead>
<tr>
<th>Game</th>
<th>State-Space Complexity</th>
<th>Game-Tree Complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nine Men’s Morris</td>
<td>$10^{10}$</td>
<td>$10^{50}$</td>
</tr>
<tr>
<td>Checkers</td>
<td>$10^{18}$</td>
<td>$10^{31}$</td>
</tr>
<tr>
<td>Othello</td>
<td>$10^{28}$</td>
<td>$10^{58}$</td>
</tr>
<tr>
<td>Chess</td>
<td>$10^{46}$</td>
<td>$10^{123}$</td>
</tr>
<tr>
<td>Risk (200 armies)</td>
<td>$10^{47}$</td>
<td>$10^{2350}$</td>
</tr>
<tr>
<td>Shogi</td>
<td>$10^{71}$</td>
<td>$10^{226}$</td>
</tr>
<tr>
<td>Risk (1000 armies)</td>
<td>$10^{78}$</td>
<td>$10^{2350}$</td>
</tr>
<tr>
<td>Go (19 x 19)</td>
<td>$10^{172}$</td>
<td>$10^{360}$</td>
</tr>
<tr>
<td>Risk</td>
<td>$\infty$</td>
<td>$10^{2350}$</td>
</tr>
</tbody>
</table>
Agenda

• Risk
• Risk Framework
• Basic Evaluation Player (BEP)
• Enhanced Evaluation Player (EEP)
• Learning Player (LP)
• Human Opponents
• Conclusion
Risk Framework – Architecture

- Player
  - RandomPlayer
  - HumanTextPlayer
- GameManager
- Rules
- Map
- Gameboard
Agenda

- Risk
- Risk Framework
- Basic Evaluation Player (BEP)
- Enhanced Evaluation Player (EEP)
- Learning Player (LP)
- Human Opponents
- Conclusion
BEP – Architecture

```
TradeFeature 1 --> TradeEvaluationFunction --> Evaluator
                  
EvaluationFunction

Feature 1 --> EvaluationPlayer --> Evaluator

TradeEvaluator
ReinforcementEvaluator
BattleEvaluator
AttackEvaluator
MoveEvaluator
```

An Intelligent Artificial Player for the Game of Risk
Michael Wolf
BEP – Decision Making Process
Agenda

• Risk
• Risk Framework
• Basic Evaluation Player (BEP)
  • Enhanced Evaluation Player (EEP)
• Learning Player (LP)
• Human Opponents
• Conclusion
EEP – Enhancements

• **Motivation: Uncoordinated Decisions of the BEP**

• **Target Continent (TC)**
  – Two new Features

• **Plans**
  – Player Elimination Plan (PEP)
  – Australia Plan (AP)
  – Continent Conquering Plan (CCP)

• **Reinforcement Distribution**
  – Several Variations
EEP – Architecture

Diagram showing the architecture of the EEP system, including TradeFeature 1 to TradeFeature n, TradeEvaluationFunction, EvaluationPlayer, Plan, PlayerEliminationPlan, ContinentConqueringPlan, AustraliaPlan, Evaluator, ReinforcementEvaluator, BattleEvaluator, AttackEvaluator, MoveEvaluator.
Rating System

• **Benchmark Player (BP)**
  - EEP with One-Max Reinforcement Distribution

• **Player Rating (PR)**
  - Four Players, One Test Player, Three BPs
  - Games Won (GW)
  - Games Played (GP)

\[ PR = \frac{GW}{GP} \times 4 \]
EEP – Enhancement Results

Player Rating

The chart shows the player ratings for different enhancements:
- BEP
- BEP + TC
- BEP + PEP
- BEP + AP
- BEP + CCP
- BEP + all Plans
- BEP + O-Max
- BEP + H-Q-Max
- BEP + Q-Q-Max
- BP

The ratings are as follows:
- BEP: 1.2
- BEP + TC: 22.4
- BEP + PEP: 78.2
- BEP + AP: 74.8
- BEP + CCP: 77.0
- BEP + all Plans: 80.4
- BEP + O-Max: 3.6
- BEP + H-Q-Max: 2.4
- BEP + Q-Q-Max: 2.0
- BP: 0.4
- Maximum: 100.0
Agenda

• Risk
• Risk Framework
• Basic Evaluation Player (BEP)
• Enhanced Evaluation Player (EEP)
• Learning Player (LP)
• Human Opponents
• Conclusion
Learning Player – Overview

• Temporal Difference (TD) Learning
  – Learning a Guess from a Guess (Bootstrapping)
  – Learning while Playing
  – Changing previous Estimates towards the current One

• TD(λ) Learning Algorithm

\[ w_{t+1} = w_t + \alpha \times (P_{t+1} - P_t) \times \sum_{k=1}^{t} \lambda^{t-k} \Delta_w P_k \]
Learning Player – Learning Risk

• Learning the weights of the linear EF

\[ w_{i,t+1} = w_{i,t} + \alpha \times (F(x_{t+1}) - F(x_t)) \times \sum_{k=1}^{t} \lambda^{t-k} f_i(x_k) \]

• Reinforcement
  – Zero or One

• Learning Steps
  – Once every Game Turn

• Normalization
Learning Player – Experiment

• Four Alpha Functions, Four LPs Each
• 10,000 Training Games
  – Six Players, Four LPs with equal Alpha, Two BPs
• Rating every 1,000 Training Games, Averaging

<table>
<thead>
<tr>
<th>Player</th>
<th>$\alpha$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Learning Player Red</td>
<td>$1/[n/1,000]$</td>
<td>0.5</td>
</tr>
<tr>
<td>Learning Player Black</td>
<td>$\frac{1}{n}$</td>
<td>0.5</td>
</tr>
<tr>
<td>Learning Player Yellow</td>
<td>$0.25$</td>
<td>0.5</td>
</tr>
<tr>
<td>Learning Player Blue</td>
<td>if $n &lt; 2,000$: $1 - 0.00025 \times n$ else: $1/[n/2,000]$</td>
<td>0.5</td>
</tr>
</tbody>
</table>
Learning Player – Alpha Functions

![Graph showing the learning progression of different players over games. The graph plots the Alpha values against the number of games played. Each color represents a different player: Black, Red, Yellow, and Blue. The graph shows how the Alpha values decrease over time, indicating learning and improvement.](image-url)
Learning Player – Results

- Red-Average
- Black-Average
- Yellow-Average
- Blue-Average
- BM-Player

Player Rating vs Training Games (in Thousands)
Agenda

- Risk
- Risk Framework
- Basic Evaluation Player (BEP)
- Enhanced Evaluation Player (EEP)
- Learning Player (LP)
- Human Opponents
- Conclusion
Human Opponents – Results

Player Rating

- Worst Human: 44.4
- Best Human: 250.0
- All Beginners: 66.7
- All Advanced: 164.7
- All Humans: 133.3
- Benchmark Player: 100.0
Agenda

• Risk
• Risk Framework
• Basic Evaluation Player (BEP)
• Enhanced Evaluation Player (EEP)
• Learning Player (LP)
• Human Opponents
• Conclusion
Conclusion

• Risk is very Complex
• Coordination of Moves is Crucial
• EEP is remarkably Strong
  – Having just three predefined Plans
• TD Learning seems Promising

➡ Learn Dynamic Plan Creation
Questions?
Thank You
Erfahrungen

- **Zeitaufwand**
  - Geschätzt
    - **Praxis**
    - **Schreiben**
  - Real
    - **Praxis**
    - **Schreiben**

- Testen/Kommentieren
- Notizen Machen
- Schreiben Parallel zur Praxis
- LaTeX