ChatGPT

Reinforcement Learning for Board Game AI: A Comprehensive Overview

Introduction

Board games have long served as benchmarks and inspiration in artificial intelligence research. In recent years, **Reinforcement Learning (RL)** – where agents learn optimal behavior through trial-and-error feedback – has achieved remarkable success in games. Classic games like chess and Go, as well as modern tabletop games, are now fertile grounds for RL research. Unlike video games or simulators, board games typically have well-defined rules and discrete turn-based structure, making them attractive testbeds. However, many strategic board games introduce unique challenges such as *hidden information*, *stochastic events* (dice rolls, card draws), and *multiple cooperating or competing players*. This report reviews the state of the art in using RL to train AI agents for board games, with emphasis on modern strategy games like **Settlers of Catan**, **Dominion**, **Small World**, **7 Wonders**, **King of Tokyo**, and related titles. We begin with a brief history of reinforcement learning and its milestones in games, then survey advances in applying RL to complex board games – particularly **deck-building games** (a genre exemplified by Dominion). Major technical approaches and milestones are highlighted, drawing on academic literature, industry projects, and open research challenges.

(We preserve citations in the format 【†】 for all referenced sources. Images are included to illustrate notable systems and results.)

A Brief History of Reinforcement Learning in Games

Early Beginnings: The concept of reinforcement-based learning traces back to the 1950s. An early landmark was Arthur **Samuel's checkers-playing program** (1959), which was the first self-learning game AI. Samuel's program improved by playing millions of games against itself and adjusting its strategy based on outcomes. Notably, after extensive self-play, it managed to defeat a human checkers master in 1962, demonstrating the potential of machine learning without explicit programming. This pioneering work – in which the program updated move values from wins/losses – effectively introduced the term *"machine learning"* and laid groundwork for later RL research.

Temporal-Difference Learning and Backgammon: A major breakthrough in modern RL was Gerald **Tesauro's TD-Gammon** in the early 1990s. TD-Gammon combined a neural network with *temporal-difference learning* (an RL algorithm) to learn the game of backgammon through self-play. With minimal built-in knowledge, the system played **1.5 million self-play games** and reached near world-champion level. By 1993, TD-Gammon's skill was just slightly below the best human players, and in a 1998 exhibition it narrowly lost to the human world champion by only 8 points over 100 games. Remarkably, TD-Gammon discovered novel backgammon strategies that expert humans adopted. This was one of the first cases of *superhuman* (or at least human-comparable) performance via RL and is frequently cited as an early success of combining RL with neural networks.

RL in Classic Board Games: Despite TD-Gammon's success, classical board games like chess and Go remained challenging. Traditional chess programs in the 1990s (e.g. IBM's *Deep Blue*) used brute-force search with expert heuristics rather than RL. It was not until the 2010s – with the advent of **deep reinforcement learning** – that RL conquered these games. A watershed moment came when **DeepMind's AlphaGo** (2016) defeated a world champion Go player. AlphaGo's design combined deep neural networks (to evaluate positions and suggest moves) with Monte Carlo tree search, and it employed both supervised learning on human games and self-play RL. The result was historic: AlphaGo achieved a **99.8% win rate** against other Go programs and beat the European Go champion 5–0, a feat previously thought to be at least a decade away. Soon after, the improved **AlphaGo Zero** version dispensed with human data entirely and learned Go from scratch via RL, surpassing the original AlphaGo.

Deep RL progress accelerated. By 2017, the generalized **AlphaZero** algorithm demonstrated that tabula-rasa self-play RL can attain superhuman skill in multiple games. Starting from random play and given only the game rules, AlphaZero achieved **superhuman performance in chess, shogi, and Go within 24 hours of training**, decisively defeating the strongest specialized programs in each game. This was a stunning validation of self-play reinforcement learning: through repeated simulation and learning, an agent could master complex perfect-information games without any handcrafted strategies. Subsequently, DeepMind's **MuZero** (2019) took a further leap by learning *without even being told the rules*. MuZero learns a model of the game dynamics on its own, focusing on aspects relevant for planning, and combines this with AlphaZero's search. Impressively, MuZero matched AlphaZero's level on Go, chess and shogi and also set new records on Atari video games. These milestones illustrate the evolution of RL from early adaptive programs to general game-playing agents that can learn any board game given sufficient computing resources.

Beyond Two-Player Classics: Reinforcement learning research has since broadened to games beyond the classical two-player, deterministic, perfect-information category. For example, **Texas Hold'em poker** (a stochastic, partially observable card game) was essentially solved at a superhuman level by AI around 2017, using variants of RL and search in combination (e.g. DeepStack and Libratus, though these relied on game-theoretic algorithms in addition to RL). In the realm of modern board games, researchers began tackling titles with **more than two players, hidden information, and complex stochastic elements** – domains often considered far more difficult for RL algorithms. We now turn to the progress in these complex board games, focusing especially on strategy games popular in hobbyist circles (sometimes called "Eurogames") and the emerging subgenre of deck-building games.

RL in Modern Strategy Board Games

Deck-Building Games: *Dominion* and Beyond

One subgenre of board games that has garnered attention is **deck-building games**, where players gradually build a personal deck of cards from a common card supply. The flagship example is **Dominion** (2008), a seminal deck-building game that inspired many others. In Dominion, each game starts with a random set of kingdom cards available for purchase, meaning *no two games are alike*. This variability poses a unique challenge for AI: the agent must adapt its strategy to different card combinations on the fly. Dominion also involves hidden information (each player's shuffled deck and hand) and randomness (shuffling, drawing cards), making it a complex environment for RL. Despite these difficulties, Dominion has recently emerged as a "new frontier" for AI research [1], and several research teams and companies have attempted to build Dominion-playing agents.

Early Approaches: Initial attempts to create Dominion AIs did not use deep RL, relying instead on heuristics or evolutionary algorithms. For instance, **Provincial** (2014) was an AI that used **genetic algorithms and competitive co-evolution** to evolve effective buying strategies for a given set of kingdom cards. Provincial would simulate millions of games for each new kingdom to discover strong strategies [2]. It was notably strong – competitive with experienced human players – but had limitations. Because it trained separately for each kingdom setup, it lacked a general policy and did not leverage online learning during play. Other early bots (e.g. those in the online platform "Dominion Online") were largely rule-based, using hardcoded policies for what cards to buy or play, which made them brittle when encountering new or unanticipated card interactions.

Reinforcement Learning in Dominion: In the last few years, researchers have applied modern RL algorithms to Dominion, with promising results. One of the first such efforts was a 2018 project by Ian Davis, who trained agents on a simplified Dominion setup using policy-gradient methods [3]. Davis's agents learned reasonable purchasing strategies on a limited card set (the basic treasures and victory cards), though extending to the full game remained challenging. Around the same time, a Stanford class project (2018) used deep reinforcement learning on Dominion's base set and reported beating some of the built-in bots [3]. These early RL agents, however, were restricted to very small subsets of Dominion due to the enormous state and action space of the full game.

A significant academic milestone came in 2023, when Gerigk and Engels introduced a deep RL agent for Dominion that leveraged **geometric state representations and a Soft Actor-Critic algorithm**. This agent was able to learn effective play styles, outperforming previous approaches. The authors emphasized that deck-building games like Dominion remained *"an unsolved challenge for game AI research"*, underlining the novelty of their contribution. Another 2024 study by Halawi *et al.* argued for Dominion as an AI benchmark and provided an RL **baseline bot** that learned via self-play to beat most existing Dominion bots [1]. Their **DQN-based agent** (using a Rainbow DQN algorithm) achieved particularly notable performance: it handily beat simple heuristic bots and even **won or tied against "Provincial" about 2/3 of the time**. This was a striking result, since Provincial had been the strongest known AI previously. The DQN agent did struggle against one specific type of strategy (a particular "engine" combo) due to insufficient exploration of that tactic during training, highlighting how diverse Dominion's strategy space is. Nonetheless, demonstrating that an RL agent can **"crush existing bots"** in Dominion is a major step forward.

Industry Application – *Neural Net Dominion AI*: Parallel to academic research, industry developers have also tackled Dominion using RL. **Temple Gates Games**, a digital board game studio, obtained the license to develop an official Dominion app and created a state-of-the-art Dominion AI. Their approach, as described by Temple Gates developers and summarized by AI researcher Alex Irpan, was inspired by AlphaZero. The Dominion AI uses a **Transformer-based neural network** (to accommodate variable card sets and game history) as both a policy and value function, and it is trained via self-play combined with **Monte Carlo Tree Search (MCTS)** for decision-time lookahead. In essence, the agent evaluates moves with its neural net and uses MCTS to explore a few turns ahead, then learns from the self-play games, mirroring the AlphaZero template. Notably, the developers did **not** inject any Dominion-specific expert rules; the AI learns purely from gameplay and outcomes. This is the first Dominion AI to operate at a fine-grained decision level (choosing how to play each card and buy each turn) without relying on human-crafted heuristics. According to reports, the Temple Gates bot is already quite strong, especially at classic "Big Money" strategies (simple tactics focusing on treasure and Province buying). The developers gradually train it on more card expansions, resetting training as new cards are added, effectively *"increasing the difficulty"* over time. While its ultimate skill against top human players remains to be seen, this effort underscores that **AlphaZero-**

style reinforcement learning can be applied to a complex, random, multiplayer game like Dominion. Researchers had speculated that pure RL would learn too slowly due to Dominion's high variance and randomness, but combining RL with MCTS and modern neural nets appears to overcome many of these issues.

Technical Challenges in Dominion: Why is Dominion particularly hard for RL? First, the **action space** is huge and *state-dependent*. At any point, a player's legal moves depend on their hand, available buys, remaining supply cards, etc. Handling a varying action set (e.g. only legal card purchases) was non-trivial; early implementations had to mask invalid actions to stabilize learning. Second, the game is partially observable (opponents' decks and hands are hidden), and outcomes have a long horizon – victory is determined only after dozens of turns, making the reward signal sparse. Approaches like giving small intermediate rewards (e.g. a bonus for buying a victory card) were tried to guide learning. Third, the strategy space is extremely rich: because the set of kingdom cards changes every game, a good agent must generalize across many possible game configurations. This is why Dominion is proposed as a next-generation benchmark for RL – it requires **adaptive, generalist play** rather than optimizing a fixed strategy. The Dominion Online community has also provided a boon in the form of data: a dataset of *2 million human games* was compiled as a potential learning resource [4]. Future research may incorporate *imitation learning* from this data to give agents a head-start, similar to how AlphaGo was pre-trained on human expert moves. Indeed, using human data or domain knowledge can greatly speed up learning in complex games – something the next case study demonstrates.

Case Study: *7 Wonders Duel* – AlphaZero Meets Card Drafting

7 Wonders Duel is a popular 2-player strategy board game (a spinoff of the multiplayer game *7 Wonders*). It features an intricate card-drafting system, multiple victory conditions, and some hidden information (cards are revealed gradually). In 2024, Paolini *et al.* introduced **ZeusAI**, an AlphaZero-inspired RL agent for 7 Wonders Duel. ZeusAI achieved a milestone: it learned to play at the level of top human experts *without any human supervision*, purely via self-play reinforcement learning. Using a combination of **MCTS and a Transformer neural network**, ZeusAI was trained from scratch on the full game rules. The choice of a Transformer architecture (as opposed to the convolutional networks used in AlphaZero) was motivated by the complex, non-spatial state of the game (various cards, tokens, and effects). Additionally, because 7 Wonders Duel has some randomness (card shuffles) and non-deterministic effects, ZeusAI included a mechanism to limit random rollouts in MCTS for efficiency.

The results were impressive. In tests against top human competitors, **ZeusAI won 26 out of 38 games (~68% win rate)**, a performance that potentially exceeds human grandmasters in this game. The AI not only rediscovered known strong strategies, but also made "counter-intuitive" moves that surprised expert players – suggesting it found novel tactics in a game that humans have studied for years. The authors even used the trained agent to evaluate game balance; for example, they measured the advantage of going first (found to be ~66.8% win rate for the first player) and tested rule modifications that could level the field. This illustrates how a superhuman RL agent can serve as a tool for game *analysis and design*, not just competition. The ZeusAI project is one of the clearest demonstrations that **AlphaZero-like reinforcement learning can handle a modern tabletop game with complex rules and multiple win conditions, achieving expert-level play**. It validates the approach of combining self-play RL with search and state-of-the-art neural architectures on non-traditional board games.

Multi-Player and Stochastic Games: *Settlers of Catan* and Others

Many modern board games involve more than two players, significant randomness, and elements of negotiation or cooperation. **Settlers of Catan** (simply *Catan* in recent editions) is an iconic example: a 4-player economic game with dice rolls, hidden resource cards, and player trading. These features pose a substantial challenge for RL algorithms, which were originally developed and tested mostly on two-player zero-sum games. Nonetheless, several researchers have ventured into Catan AI, making incremental progress:

- **Early Research:** One of the first learning-based Catan AIs was by Pfeiffer (2004), who applied reinforcement learning to certain strategic decisions, but this approach required a lot of built-in human knowledge and heuristics to be effective. In effect, it was a *hybrid* of RL with manually scripted strategy, and it did not tackle the full game in a general way.

- **Deep RL Attempts:** More recently, **Gendre and Kaneko (2020)** introduced a deep reinforcement learning agent for a simplified two-player version of Settlers of Catan. They limited the game to 1v1 with no trading (since trading is inherently a 3+ player interaction) and used a novel neural network architecture (a "cross-dimensional" network) to handle the complex state inputs (board, cards, etc.). The notable result here was that *for the first time, an RL agent outperformed the best built-in heuristic AI (the "jSettlers" bot)*. In other words, their self-play trained agent learned to play better than the strongest scripted Catan AI for that two-player scenario. This was an encouraging proof-of-concept that deep RL *can* learn meaningful strategies in Catan. However, the restriction to two players without trade meant many of Catan's hardest aspects (multi-agent cooperation and negotiation) were not addressed.

- **Full 4-Player Catan:** The real goal is to train agents for the full 4-player game including trading. This is exceedingly difficult because the action space is enormous (players can propose arbitrary trades, build in many locations, play development cards, etc.) and the environment is stochastic and partially observed. In 2021, researcher **Henry Charlesworth** undertook this challenge as a personal project, building a custom Catan simulator and applying deep RL (Proximal Policy Optimization, PPO) to train agents. Over roughly a month of training (about 450 million decision steps), his agents did improve substantially – demonstrating non-trivial strategies – but they still fell short of human skill. For example, the agents learned sensible early-game moves like placing initial settlements on strong resource spots and avoiding obvious pitfalls. They also learned to avoid hoarding too many cards (to minimize losses when a 7 is rolled, which forces discards). These behaviors indicate the agents grasped some strategic fundamentals of Catan. However, other aspects remained poor: notably, they **did not master trading**. The RL agents would frequently propose worthless trades that no rational player would accept, and sometimes even accept disadvantageous trades. This suggests that the credit assignment for negotiating in a multi-agent setting is extremely hard – the agents had little incentive to propose fair trades when self-play partners occasionally accepted unfair ones, reinforcing bad behavior. Charlesworth reports that disabling trading entirely leads to more sensible gameplay, whereas with trading turned on, the training produced a lot of "spam" proposals that human players would find tiresome. His conclusion was that, with the compute and algorithms used, the agents *"definitely got better over time, but [are] still not close to the standard of a good human player"*. This is not surprising given that Catan's state-action space is far more complex than games like Go or chess, and current deep RL methods are notoriously sample-inefficient (requiring astronomical numbers of training games) and often struggle with the non-stationarity of multi-agent learning.

Progress and Prospects in Catan: Despite the challenges, each attempt at Catan AI has pushed the envelope. Research has identified some solutions to subproblems (e.g. better state representation for mixed discrete and geometric features, or framing trade proposals as a separate action head). There have also been hybrid approaches, such as Xenou *et al.* (2014) who trained a DQN agent that only handled the trade negotiation part, while using an existing AI (jSettlers) to handle all other decisions. This modular approach showed that an RL agent could learn *when to accept or reject trades* effectively, if the rest of the gameplay was fixed. Going forward, applying **multi-agent reinforcement learning (MARL)** techniques and incorporating *human data* could accelerate learning for games like Catan. For instance, one could train agents by first imitating human trading behavior (from recorded games) and then fine-tuning with self-play RL – akin to how **Meta's CICERO** AI approached the game **Diplomacy** (more on this below). Indeed, Catan shares some properties with Diplomacy in terms of requiring negotiation and alliance formation, albeit in a simpler form.

Other Games (Small World, King of Tokyo, etc.): The user specifically inquired about games like *Small World*, *7 Wonders*, *King of Tokyo*, and similar titles. Direct research on these specific games is relatively sparse compared to Go, Poker, or even Catan/Dominion. However, general techniques developed for one game often transfer to others in the same family:

- **Small World** is a territory-control game for 2-5 players with deterministic combat but random pairings of special abilities. We did not find notable published RL research on Small World, likely because its state space (a changing map and combinations of powers) is complex and there hasn't been an established research environment for it. It remains an open domain where an AlphaZero-like or MARL approach could be tried in the future.

- **King of Tokyo** is a light dice and press-your-luck game. Its simpler mechanics (Yahtzee-style dice rolls, up to 6 players) might actually make it easier for a basic Q-learning or MCTS agent to handle, but we found no public research dedicated to it. A well-designed simulator plus self-play could probably yield a reasonable King of Tokyo bot, but the game's strategic depth is not as rich as others, so it may be of less academic interest.

- **7 Wonders (multiplayer)**: Apart from the 2-player Duel variant discussed earlier, the original 7 Wonders (3–7 players, simultaneous card drafting) presents a challenge due to simultaneous actions and partial information (you only see your neighbors' plays fully). Some work in general card drafting AI exists (e.g. for drafting in card games like Magic: The Gathering or Hearthstone arena), often using simulation and heuristic rollouts rather than pure RL. An RL agent for 7 Wonders would need to deal with a large joint action space and could perhaps use methods for *impartial games* or multiplayer Nash equilibria. As of this writing, we have not seen a dedicated 7 Wonders RL publication, but the success on 7 Wonders Duel suggests that tackling the multiplayer version is within reach by extending the methods to simultaneous decision-making (perhaps using an iterative best-response or policy gradient approach in a multi-agent setting).

- **Other Deck-Builders:** Aside from Dominion, other deck-building games (e.g. *Ascension*, *Star Realms*, or digital ones like *Hearthstone* and *Slay the Spire*) have seen some AI research. A 2016 study by García-Sánchez *et al.* applied **evolutionary algorithms** to the deck-building aspect of the digital card game Hearthstone [5]. They evolved effective decks and policies for playing cards, demonstrating the usefulness of evolutionary search in the deck construction phase [5]. On the RL side, Hearthstone proved very difficult due to enormous state space and hidden information, but a few works (e.g. DeepMind's experiments with **AlphaZero-style methods on Hearthstone Battlegrounds**) indicate growing interest. Furthermore, a recent AI competition around the deck-building game **"Tales of Tribute"** (from *Elder Scrolls Online*) provided a simulator and challenge for RL agents. This shows that the research community is actively seeking generalizable methods for *deck-building strategy optimization*. Techniques like policy gradient and tree search are being adapted to handle the combinatorial explosion of possibilities that deck-building entails (e.g. which card to buy at each opportunity, which introduces long-term ramifications on future draws).

Multi-Agent and Cooperative Games: Communication and Negotiation

When board games include **three or more players**, the competitive dynamics change fundamentally. Self-play between two agents (as used in AlphaZero) no longer directly applies, because with more players the game is not zero-sum and learning an optimal strategy is more complex (the environment is non-stationary as other learning agents adapt). Researchers are exploring *multi-agent reinforcement learning* (MARL) in such contexts. A standout example is **Diplomacy**, a classic 7-player board game of alliances, negotiation, and betrayal. In 2022, Meta AI introduced **CICERO**, an agent that combines RL with natural language processing to play online Diplomacy. CICERO first learned to negotiate by analyzing human dialogue data, then used RL planning to propose actions and messages during the game. In a league of human Diplomacy players, CICERO achieved **double the average score of human players and ranked in the top 10%**, making it the first AI to reach human-level performance in a social negotiation setting. CICERO's success is a milestone not just for games but for AI in multi-agent environments generally – it required reasoning about other players' beliefs and intentions, a kind of theory-of-mind modeling.

While Diplomacy is not one of the originally listed games, its inclusion here is relevant: it demonstrates how far multi-agent RL has progressed. Techniques used by CICERO (such as training with a mixture of imitation learning from human game logs and self-play RL, and using dialogue models to negotiate) could potentially be adapted to games like Catan or others involving trading and cooperation. Similarly, **cooperative board games** (where all players work against the game) have been studied through the lens of MARL. A prominent case is the card game **Hanabi**, which is fully cooperative and involves implicit communication through actions. The *Hanabi Challenge* was proposed as a benchmark for AI research on collaborative, imperfect-information games [6]. Deep RL algorithms have been applied to Hanabi to learn communication protocols, with some success (agents can achieve near-perfect scores in simplified settings), though aligning AI communication with human-understandable strategies remains an open challenge.

In summary, RL in multi-player board games is still in its infancy relative to two-player games. Each added player (or teammate) introduces new complexity in the learning problem. However, the field is rapidly advancing: recent successes like **Cicero in Diplomacy** and **ZeusAI in 7 Wonders Duel** show that even games once deemed too complex for AI are yielding to creative combinations of deep learning, game-theoretic reasoning, and massive compute.

Technical Approaches and Key Insights

Across the efforts discussed, several **technical themes** emerge:

- **Self-Play and League Training:** For competitive games, **self-play reinforcement learning** has proven to be a powerful paradigm (AlphaGo, AlphaZero, etc.). By playing against copies of itself, an agent can iteratively bootstrap its skill. In multi-player games, self-play can be augmented by training against a *population or league* of agents to stabilize learning (preventing cyclic dynamics). For example, in Dominion and Catan experiments, researchers often maintain a set of past agent versions and pit current agents against older ones to ensure continual progress. This is analogous to the *"league training"* used in AlphaStar (for StarCraft II) and OpenAI Five (for Dota2).

- **Monte Carlo Tree Search (MCTS):** Combining learning with search at decision time can greatly enhance performance, especially in games with high branching factors. The AlphaZero algorithm's integration of MCTS was key to its success. We see this carried over to Dominion (Temple Gates AI uses an MCTS guided by the network) and 7 Wonders Duel (ZeusAI's MCTS over game states). MCTS helps mitigate the depth of the credit assignment problem by effectively looking ahead a few moves and providing a higher-quality training signal to the network.

- **Neural Network Architecture:** Early game RL systems like TD-Gammon used shallow neural networks. Today's agents use deep architectures tailored to the game's state representation. Convolutional Neural Networks (CNNs) are natural for grid-based games (chess, Go). For card-based or more abstract games, **Transformers** have gained popularity because they can flexibly encode sets of features (e.g. cards in hand, cards in supply) without assuming a particular spatial structure. In Dominion, a Transformer was used to encode game states due to the varying card sets and to capture longer-term dependencies (such as what cards were bought earlier). Likewise, in 7 Wonders Duel, the Transformer outperformed a CNN because the game state isn't a fixed image – it's a collection of cards, tokens, and tracks. Designing the input features is also crucial: researchers often encode state in terms of **inventories** (how many of each resource/card a player has) and **observable actions** (e.g. binary flags for which moves are legal). Action masking (ensuring the network can only choose legal moves) was necessary in games like Dominion and Catan to avoid meaningless outputs that hinder learning.

- **Reward Shaping and Intermediate Objectives:** In pure self-play, the reward is typically win=1, loss=0 (or -1), and perhaps draw=0.5. This sparse reward can make learning slow. Thus, researchers sometimes add intermediate rewards or curriculum learning. For example, an agent might get a small reward for *collecting certain cards* or *reaching a score threshold* to guide it initially. In Dominion, one experiment gave a tiny reward for buying a Victory card to encourage the agent to value points. Care must be taken with shaping to not alter the optimal strategy (one must avoid agents optimizing the shaped reward at the expense of actual winning). An alternative is **curriculum learning**: start the agent on easier versions of the game (e.g. Dominion with a very small set of cards, or Catan against a fixed strategy opponent) and gradually increase the difficulty as it improves. Temple Gates effectively did this by training on base Dominion, then introducing expansions progressively, rather than learning all 300+ cards at once.

- **Combining Learning Paradigms:** In complex games, a pure RL approach from scratch can be prohibitively slow. Successful projects often mix methods. CICERO (Diplomacy) combined *behavioral* *cloning* from human game data with RL fine-tuning. AlphaGo combined supervised learning from human moves with RL. Even AlphaZero, which did not use external data, relied on heavy use of MCTS at training time to focus the learning updates on promising moves. In academic research on board games, one sees hybrids like **Evolutionary Reinforcement Learning** (using genetic algorithms to evolve policies, with reinforcement signals as part of fitness) or **Expert Iteration**, which alternates a planning (expert) phase with a learning (student) phase. The Provincial Dominion AI, for example, can be seen as evolutionary search that finds strong strategies, which could then be used to supervise a policy network. Indeed, a recent trend is to use *search or simulation data to train neural nets*, effectively distilling the strengths of both search-based and learned approaches.

- **Generalization and Inductive Biases:** A key issue for board game RL is how well an agent trained on certain scenarios can generalize to new ones. Dominion exemplifies this because of its random card sets. One agent, as noted, was trained to beat specific "kingdoms" but failed on unseen ones. Modern approaches aim for general strategies that work across a distribution of games. Designing state features that capture abstract properties (instead of overfitting to particular cards or board layouts) helps. For instance, encoding a Dominion card by its effects (draws cards, gives coins, etc.) could let a network handle new cards with similar effects. Research in *meta-RL* and *procedurally generated environments* is relevant here: an agent can be trained on many random variations so that it learns to adapt to any given instance. Open frameworks like **OpenSpiel** and **RLCard** provide a library of games for exactly this purpose, allowing researchers to train agents that learn the *rules* of new games on the fly. While OpenSpiel includes mostly simpler games (e.g. Tic-tac-toe, Go, Poker), the philosophy is to encourage algorithms that are not single-game specialist but can handle a variety of games, much like a human can learn many board games.

Conclusion and Outlook

The intersection of reinforcement learning and board games has evolved from early experiments in checkers to superhuman play in some of the most challenging games known. We have witnessed **major milestones**: Tesauro's TD-Gammon in the 90s showed self-play RL could rival pros in backgammon; AlphaGo and AlphaZero (2016–2018) then decisively beat human champions in Go, chess, and shogi. These successes mostly involved two-player perfect-information games. Today, the frontier has shifted to more complex, **multi-faceted board games**. Dominion, as a **deck-building game with high randomness and variety**, is emerging as a benchmark for RL, with baseline agents now surpassing the best heuristic bots. Multiplayer games like Catan remain tough nuts to crack, but gradual progress (from 2-player subsets to partial-function agents to full-game prototypes) suggests that stronger models and more compute can eventually yield agents that compete with humans. Notably, in games incorporating elements of *negotiation or cooperation*, AI has begun to excel – the Diplomacy-playing CICERO being a prime example of combining language and RL to interact with human players on their terms.

For the specific case of **Dominion AI**, the current state of research is very encouraging. Multiple independent efforts (academic and commercial) have demonstrated that deep RL agents *can* learn effective Dominion strategies. The use of techniques like MCTS and transformers has overcome many initial hurdles (such as high variance and large state space). A 2024 paper even provides an **open dataset of 2M human Dominion games and an open-source Dominion simulator with an RL agent** [4], which will spur further research and competition. The *Dominion AI Project* for which this report is compiled stands to benefit from all these developments. By leveraging established algorithms (like PPO, DQN, or AlphaZero-style self-play) and the lessons learned (e.g. action masking, curriculum training, integrating search), one can build upon prior agents to push Dominion AI to new heights. A remaining challenge is achieving **expert-level play across all Dominion expansions** – a vast space that no published AI has fully conquered yet. However, with techniques such as progressive training (introducing expansions one by one) and perhaps utilizing the trove of human gameplay data for pre-training, this goal appears on the horizon.

Finally, it's important to note that **performance benchmarks**, while useful, are not the sole focus of research now. Equally important is *interpreting* the strategies learned by RL agents (are they reasonable, novel, exploitable?), ensuring agents can handle the **variance** inherent in games (Dominion's shuffle luck or Catan's dice luck), and extending these methods to real-world problems. Board games are idealized models of strategic decision-making; insights from board game RL are beginning to inform economics, logistics, and beyond. The rich strategies in games like Dominion – which involve **long-term planning, adaptation to random events, and even a bit of bluffing** – mirror problems in business and finance. In this sense, achieving superhuman board game bots is not just a stunt; it is a step toward AI that can handle complex sequential decision tasks under uncertainty.

In conclusion, reinforcement learning has achieved remarkable successes in board game AI, from classic games to modern complex ones. **Deck-building games** have gone from being "too hard for AI" to a proving ground for new techniques, with Dominion now positioned as a benchmark for generalizable, robust RL algorithms [1]. **Multi-player strategy games** are increasingly in focus, and though challenges remain, each breakthrough (like 7 Wonders Duel's expert-level agent or Diplomacy's human-level negotiator) expands the frontier of what AI can do. The journey of training bots to play board games is far from over – but with each innovation in deep RL, the once insurmountable games are falling, one by one, to the prowess of learning algorithms. The coming years may well see AI champions in games like Catan, or even in entire genres of games, trained by powerful reinforcement learning systems that build upon the milestones reviewed in this report.

Sources:

- Samuel, A. (1959). *Some Studies in Machine Learning Using the Game of Checkers*. IBM J. Res. Dev. (Early checkers self-learning program).
- Tesauro, G. (1992-1995). *Temporal Difference Learning and TD-Gammon*. **(RL backgammon agent approaching top human level)**.
- Silver, D. *et al.* (2016). *Mastering the Game of Go with Deep Neural Networks and Tree Search*. Nature. **(AlphaGo, first superhuman Go AI)**.
- Silver, D. *et al.* (2018). *A General Reinforcement Learning Algorithm that Masters Chess, Shogi, and Go through Self-Play*. Science. **(AlphaZero, tabula rasa superhuman play in multiple games)**.
- Schrittwieser, J. *et al.* (2020). *Mastering Atari, Go, Chess and Shogi by Planning with a Learned Model*. Nature. **(MuZero, learning without rules)**.
- Halawi, D. *et al.* (2024). *Dominion: A New Frontier for AI Research*. arXiv preprint. **(Proposes Dominion as benchmark, RL bot beats prior AIs)** [1].
- Gerigk, J., & Engels, S. (2023). *Playing Various Strategies in Dominion with Deep Reinforcement Learning*. Proc. AAAI Conf. on AI and Interactive Digital Entertainment. **(Geometric deep RL agent for Dominion)**.
- Temple Gates Games (2021). *Dominion Neural Network AI Announcement*. **(Transformer + MCTS AlphaZero-style Dominion AI)**.
- Davis, I. W. (2018). *Training an AI for the Card Game Dominion (Blog Series)*.

- Paolini, G. *et al.* (2024). *Learning to Play 7 Wonders Duel Without Human Supervision*. arXiv preprint. **(ZeusAI, AlphaZero-like agent, ~68% win vs top humans)**.
- Charlesworth, H. (2021). *Learning to Play Settlers of Catan with Deep RL (Blog)*.
- Gendre, Q., & Kaneko, T. (2020). *Playing Catan with Cross-Dimensional Neural Networks*. IEEE Conf. on Games. **(2-player Catan RL agent beating heuristic AI)**.
- Pfeiffer, M. (2004). *Reinforcement Learning of Strategies for Settlers of Catan*. **(Early attempt with heuristics)**.
- Meta AI (2022). *Human-Level Play in the Game of Diplomacy by Combining Language Models with Strategic Reasoning*. Science. **(CICERO agent achieves top 10% performance in Diplomacy)**.
- Bard, N. *et al.* (2019). *The Hanabi Challenge: A New Frontier for AI Research*. **(Proposal for cooperative multi-agent RL benchmark)**.
- García-Sánchez, P. *et al.* (2016). *Evolutionary Deckbuilding in Hearthstone*. IEEE Conf. on Computational Intelligence and Games. **(Uses evolutionary algorithms for card deck optimization)** [5].

(Images: Settlers of Catan RL Simulator screenshot – Henry Charlesworth, 2021)

---

[1] [4] Dominion: A New Frontier for AI Research
https://arxiv.org/html/2405.06846v1

[2] [3] A New Online Dominion Client Approaches
https://www.alexirpan.com/2021/05/23/dominion-temple.html

[5] Introducing Tales of Tribute AI Competition This work was supported by the National Science Centre, Poland under project number 2021/41/B/ST6/03691.
https://arxiv.org/html/2305.08234v4

[6] [1902.00506] The Hanabi Challenge: A New Frontier for AI Research
https://arxiv.org/abs/1902.00506