
Thought Piece: Game Changer

The historic chess matches between Garry Kasparov and IBM's Deep Blue supercomputer is often thought of as pivotal moment in changing perceptions of machine intelligence.

Kasparov won the first match in 1996 with three wins, two draws and one loss. IBM returned for a rematch in 1997 with "Deeper Blue", a computer with double the processing power and improved positional analysis from a grandmaster, Joel Benjamin. Deeper Blue won two games, drew three and lost one, marking the first time a reigning world chess champion had lost to a machine.

But little known and scarcely acknowledged is that the watershed moment relevant to today's advancements in AI had taken place 5 years earlier, also at IBM, when Gerald Tesauro developed a computer program to play the game of backgammon that defeated the leading experts at the time.

The difference between the two isn't just in the game, but how the computer was programmed to play it.

Deep Blue operated as an expert system, using its immense computational power to execute a brute-force search through millions of moves, with each position's value determined by a complex evaluation function handcrafted by human grandmasters.

In contrast, TD-Gammon was not programmed with any knowledge of human tactics. It was instead exposed to millions of simulated games and allowed to play against itself, learning tactics from the final outcomes.

Deep Blue was effectively a brilliant student of human chess players whereas TD Gammon became its own backgammon teacher through trial and error.

The breakthrough in TD-Gammon that foreshadowed today's AI advancements lay in its use of parallelism. The program was trained by playing millions of simulated games, rapidly testing and refining tactics. This process demanded running many scenarios in parallel rather than sequentially.

With this context in mind, it is fitting that the most important hardware breakthroughs in modern AI came from a company built to enhance gaming: Nvidia.

Graphics Processing Units (GPUs), were designed to render video game frames by performing thousands of simultaneous calculations, proved ideal for parallel computation.

In the early 2000s, an Nvidia engineer realised that chaining multiple GPUs together could dramatically increase the performance of games like Quake III. More importantly, he saw that this same architecture, could be repurposed for general uses, unlocking entirely new possibilities in fields like scientific computing, and eventually, artificial intelligence.

In that sense, the story of modern AI doesn't begin with a line of code, but with the roll of a dice, and establishing this historic link between games and AI raises interesting questions about how strategy will develop in the future.

Although backgammon has been played for over 5,000 years, the introduction of reinforcement by machines revealed strategies that overturned centuries of conventional wisdom. TD-Gammon discovered that a widely accepted opening tactic known as *slotting*

was in fact suboptimal. As a result, strategy books written prior to 1997 had to be revised to reflect this new understanding.

The same transformation has occurred in chess. Reinforcement learning models like AlphaZero have led to the most significant evolution in chess strategy in centuries. These systems have redefined long-standing ideas around concepts like *king safety* and *piece value*, introducing more dynamic, less human-biased patterns of play. Today, with access to such machine-generated insights, a modern grandmaster like Magnus Carlsen could almost certainly defeat Deep Blue.

A similar paradigm shift unfolded in the ancient game of Go. During the 2016 match between AlphaGo and world champion Lee Sedol, the AI played a now famous “Move 37” that commentators initially believed to be a mistake. Sedol was visibly perplexed. Yet the move proved decisive, helping AlphaGo win the game.

Afterward, Sedol described it as “beautiful” and admitted: “Surely, AlphaGo is creative.” What once seemed irrational by human standards was, in fact, a higher-order insight revealed through reinforcement learning.

These breakthroughs across backgammon, chess, and Go illustrate what AI researcher Richard Sutton in 2019 called, “*The Bitter Lesson*”; the most significant advances in artificial intelligence have consistently come not from building on human knowledge, but from scaling computation and letting machines learn for themselves.

In a world where the adoption and use of AI is rising exponentially, this raises the question of what and where will the next “move 37” be?

Logically, we should expect it to emerge in real-world domains that, like games, can be distilled into environments with clear objectives, ambiguous rules, and repeatable outcomes. These are the kinds of complex systems where reinforcement learning thrives—settings that are bounded enough to simulate, yet rich enough to reward creativity.

Companies with access to vast amounts of consumer data—and the ability to segment users into distinct behavioural cohorts—are particularly well-positioned to discover patterns through large-scale experimentation.

Platforms like YouTube and Netflix have pioneered this approach, running thousands of micro-experiments in parallel to test how different types of content perform across niches. In doing so, they have often uncovered surprising insights that challenge traditional assumptions about audience preferences and content strategy.

One notable example is Netflix’s localisation strategy. In certain markets, Korean dramas dubbed in Brazilian Portuguese began outperforming expectations—not because of deliberate targeting, but due to engagement signals identified by the algorithm. As this approach matures, could Netflix or YouTube go further, using individual viewing patterns to tailor narrative structures themselves?

In the industrial and mobility sectors, the next wave of breakthroughs may come from “digital twins” and advanced simulation environments. As industries digitise, companies are increasingly able to model entire physical systems.

Waymo has logged tens of millions of real-world and simulated driving miles. With each iteration, it refines its understanding not only of driving behaviour, but of failure modes, edge cases, and human-machine interaction. As this approach gains critical scale, could it fundamentally shift our assumptions about road safety, urban design, and even the structure of insurance and regulation?

Perhaps the deeper insight from reinforcement learning is a clear demonstration that innovation emerges through repeated trial and error. The systems that surprise us do so not because they were programmed to be novel, but because they were allowed to experiment at scale and adapt beyond our biases.

For businesses, this offers a profound lesson: innovation favours those willing to embrace uncertainty and iteration. In a world increasingly shaped by algorithms that learn, the advantage will lie not with those who cling to best practices, but with those who build systems and cultures that are designed to learn.
