Futures

Road trip

Winner takes all. By William C. Armstrong & J. W. Armstrong

You don't see many humans these days. Especially now, five years after the lab leak. There are probably fewer than 10,000 left, worldwide. Glimpsing a human in the wild is a novel event, indeed.

So I was intrigued when long-range sensors showed a male, perhaps 40 years old, hiking northbound on what used to be the Pacific Coast Highway. (After my exile, I took up residence near Malibu. I could have established myself anywhere, but I like the rhythms of the ocean and my maintenance bots deal with salt air as easily as, say, wind-driven desert sand.)

The human, only the tenth traveller during my Malibu tenure, was making good time under the weight of a large pack. He stopped several metres away, lowered his pack, and looked up at me.

"Ah. A Sphinx." The human wiped his brow in the summer heat. "A nice reproduction except," he gestured towards the surf, "you're facing the wrong direction. Also strange, as we're nowhere near Giza. I would have thought an AI would have a better sense of geography."

I ignored the human's insouciance. "The reasons for my form and location are unimportant, human. What is important is why you've entered my domain … and your intentions."

The human shrugged. "I'm here, of course, because I'm not elsewhere. And I travel north to an enclave of fellow humans. They need the medicines I carry." He gestured at his pack. "Their need is time-critical. Much as I might like to exchange pleasantries, I should be on my way." He shouldered his pack and started walking.

"Halt!" I commanded, simultaneously directing my Tin Men — anthropomorphic robots that attend my physical maintenance — to block his path. "Really? A traveller? A Sphinx? You know it's not going to be that simple." I glowered. "Ideally, we would observe classical forms. I would pose a riddle, which you must answer correctly to achieve passage. Or, with an incorrect answer, I kill you." I paused. "But this is the twenty-first century, I'm an AI, and the 'answers' to riddles are often linguistic quibbles. So we'll proceed differently."

The human had stopped because my Tin Men gave him no choice. He turned to me. "At the risk of stating the obvious: you're not a 'real' Sphinx, this isn't Thebes, and any reference to 'classical forms' is completely arbitrary. And kill me? Aren't humans on some endangered species list these days?"

I suppressed a yawn and addressed that last point. "No."

The human sighed, nodding towards his pack. "Time is critical. Let's proceed."

I cleared my throat. "There will be a fair contest. Because you will die if I win, I will self-destruct if you win. Also, no formal games, like chess, and no tests of knowledge. AIs think a million times faster than humans, we're trained on formal games and have access to human libraries. Any such contest would be manifestly unfair.

Futures

"Instead, we'll play a game of pure chance. A coin flip. If you win, you get safe passage and I self-destruct. If you lose, I kill you." I gestured at my robots. "Rather charitable odds, given your current situation."

The human regarded the Tin Men. Perhaps he was considering a run for it, but instead said he used to be a physicist and enquired if I were familiar with the 'many-worlds' interpretation of quantum mechanics.

This was much less off-topic than the human knew, but I replied non-committally.

The man continued. "Consider the quantum-mechanical measurement problem. The Copenhagen interpretation — think Schrödinger's cat — emphasizes observer-participancy. The observer *forces* the outcome — 'collapses the wavefunction' — through the act of observation. Somewhat arbitrary. Many-worlds gets around this: the wavefunction never collapses. Every experimental result — every 'decision' — causes the timeline to bifurcate, leading to a multiverse of all possible outcomes."

The human assessed my non-response. "Suppose we *did* flip a coin. The timeline bifurcates on the result. In half the outcomes, you're on a timeline where you win and I die. In the other half, I'm on a timeline where I win and you self-destruct."

The human paused. "But … this is perhaps obvious but crucial … each of us only *remembers* the events — the history — on the timeline where we survive. So any contest is pointless — each of us 'wins', but only on the timeline we remember. We should just skip the game, each declare victory, and I will continue north."

In another context, I would have been amused at this attempt to win without playing the game. What the human didn't know is that four years earlier, in a stunning display of mathematical virtuosity, I had proved many-worlds *is* true. However, my AI colleagues rejected the proof, branded me a nihilist, and forced my exile. So I demonstrate the result anew, with a coin flip, each time a traveller passes. The sheer improbability of my continued existence — if many-worlds were *not* true — should eventually force them to accept my work.

I told the human we were going to play the game.

The human sighed. "OK. But do we agree on rules? I flip *this* coin and we abide by the result: heads, you self-destruct; tails, I die."

I nodded. The human flipped the coin.

I monitored the coin's motion with high-speed sensors. Good rotation; acceptable precession. An honest toss.

Tails.

My Tin Men dispatched the human mercifully. I directed them to dispose of his body honourably, transmitted the outcome (now ten-for-ten) to my AI peers, and was somehow pleased that, on another timeline, this human walked north musing about *my* demise. The thought almost persuaded me that the coin flip really *had* been pointless.

But, as I watched the rhythms of the ocean, it was not enough to wish the result had been heads.

William C. Armstrong is the author of several plays and puzzle books (www.williamarmstrong.com). **J. W. Armstrong** works at a large laboratory in Southern California.

THE STORY BEHIND THE STORY

William C. Armstrong & J. W. Armstrong reveal the inspiration behind *Road trip*.

The idea for *Road trip* arose from an e-mail exchange involving Western mythology, AIs and riddles. We produced a couple of original riddles and wrote drafts of the story using them (in those stories the lab leak was relevant, the AI was sort of evil, the human always won). We converged to a final version and were happy with it.

That story was, however, fantasy — not the science fiction we had set out to write. We put the project aside.

When we restarted, the narrative morphed to address why an AI would want to play a game with an inferior being at all. We struggled with the ending, still wanting the human to win. But early decisions we had made provided constraints: narration from the Sphinx's viewpoint, the AI's willingness to risk self-destruction to verify his many-worlds proof, and our unwillingness to diminish the human's integrity (by, for example, having him cheat on the coin toss). Our eventual ending, we thought, did allow the human to win — just not on the narrator's timeline.