AI Assistant Scientist: Aiding scientific discovery with quantum-like evolutionary algorithm

Lizhi Xin[1, †], Kevin Xin[1, *]

[1]XINVISIONQ, INC.

†Contributing author. E-mail: lizhi@xinvisionq.com;
*Corresponding author. E-mail: kevin@xinvisionq.com;

Abstract

The greatest challenge of scientific discovery is how unknown becomes known. While there have been ways set forth to go about scientific discovery, most require rigorous work done by humans, and involve human scientists trying their best to sift out the most important information from vast amounts of raw data. This paper presents an AI assistant scientist that utilizes quantum superposition principle to model many different theories and applies genetic programming to evolve the most satisfactory theory from many possible ones. Setting out from an information perspective, multiple AI agents cooperate to find the valuable information from raw data: (1) produce a function for natural phenomena with complete information obtained from experimental data, and (2) produce a matrix for natural phenomena where complete information cannot be obtained from experimental data. Using the freefall trajectory of a light sphere and Schrodinger's Cat simulated thought experiment as case studies, we show that the AI assistant scientist is able to reconstruct the past trajectory and predict the future trajectory of the light sphere with the function it produces, and to reconstruct the cats' past states and probabilistically predict the cat's future states with the matrix it produces. We believe that the key to scientific discovery is how to obtain as much valuable information from raw data as possible and this can be done by the AI Assistant scientist that's powered with the quantum-like evolutionary algorithm that we've developed.

Keywords: scientific discovery, genetic programming, machine learning, AI assistant scientist, quantum-like decision theory, quantum-like evolutionary algorithm

1 Introduction

Fundamentally the basis of scientific discovery is to find the regularities of the observable data on natural phenomena; this task has generally been accomplished by scientists. The question now becomes: is there a possibility of developing an algorithm that can automatically find regularities from raw data? Recently there has been research that has proposed a variety of algorithms to find regularities from raw data. Roger Guimera and Marta Sales-Pardo [1] developed a symbolic regression algorithm called the Bayesian machine scientist; Patrick Langley [2] developed BACON to rediscover Kepler's third law; Lipson and Michael Schmidt [3] applied genetic programming to develop an algorithm called Eureqa which successfully recover equations describing the motion of one pendulum hanging from another; Steven Brunton, Joshua Proctor, and Nathan Kutz [4] developed an algorithm by applying sparse regression; Miles Cranmer et al. [5] developed a symbolic regression algorithm directly inferring Newton's law of gravitation; Max Tegmark and Silviu-Marian Udrescu [6] developed an algorithm called "AI Feynman" to rediscover 100 equations from the Feynman Lectures on Physics; Cristina Cornelio et al. [7] developed an algorithm called "AI-Descartes" for scientific discovery.; and Krenn et al. [8] wrote a survey article about scientific understanding with artificial intelligence.

This paper proposes a quantum-like evolutionary algorithm [9-10] based on the quantum superposition principle and Darwinian natural selection that will attempt to find regularities from raw data through machine learning. Quantum superposition is utilized to model the many different “theories” to describe natural phenomena, and the evolutionary algorithm is then applied to optimize the most satisfactory “theory” from the many different ones.

(1) If complete, accurate information of a natural phenomenon has already been observed through experiments: the quantum-like evolutionary algorithm will construct a “function tree” population and by means of natural selection evolve the most satisfactory function that is a definite description of that natural phenomena. Using the freefall data of a light sphere, we illustrate how the quantum-like evolutionary algorithm is able to find the function to fit the trajectory path of the light sphere.

(2) If complete, accurate information of a natural phenomenon cannot be observed through experiments: the quantum-like evolutionary algorithm will construct a “matrix tree” population and by means of natural selection evolve the most satisfactory matrix to assess the trend of that natural phenomenon along with a probabilistic forecast. Using simulated data of Schrodinger’s Cat thought experiment, we illustrate how the quantum-like evolutionary algorithm is able to find the matrix that will assess whether the cat is dead or alive without opening the “box”.

The structure of the paper is as follows: Section 2 details the methodology. Section 3 are the results. Section 4 is the discussion.

2 Methods

Algorithm 1 Genetic Programming Algorithm

Input:

- Historical dataset $\{(t_k, x_k), k = 0, \cdots, N\}$;
- Setting:
  (1) Operation set F;
  (2) Dataset T;
  (3) Crossover Probability = 70%; Mutation probability = 5%.

Initialization:

- Population: randomly create 300 individuals.

Evolution:

- Loop: for i = 0 to 80 generations.
  a) Calculate fitness for each individual based on the historical dataset;
  b) According to the quality of fitness:
     i. Selection: selecting parents.
     ii. Crossover: generate a new offspring using the roulette algorithm based on crossover probability.
     iii. Mutation: randomly modify the parent based on mutation probability.

Output:

- The fittest individual.

Where the operation set F and dataset T will be given according to the specific case presented, as shown in Algorithm 1. The quantum-like evolutionary algorithm utilizes Genetic Programming (GP) [11-14] to optimize the "function tree" and "matrix tree", where GP is based on Darwinian evolution theory. The theory of Darwinian natural selection states that the fittest individual of a population survives by evolving generation after generation through means of crossover, mutation, and selection. What is crucial to the algorithm is to find a fitness function that will evaluate how fit each individual is in relation to the environment. The fitness function can be defined respective to the problem being presented.

2.1 Function Tree

For the freefall of a light sphere, the operation set is just a standard set of algebraic operators, while the dataset is just a time-dependent variable and including the constants of 1 to 9.

Operation set: $\mathrm{F} = \{+, -, \times, \div, \sin, \cos, \log, \exp\}$

Data set: $\mathrm{T} = \{\mathrm{t}, 1{\sim}9\}$

The fitness function of the function tree is defined by the Mean Absolute Error (MAE):

$$\mathrm{fitness}_{\mathrm{functionTree}} = \sum_{i=1}^{N} |x_i - y_i| \tag{1}$$

Where x_i is the observed value, y_i is the calculated value, N is the number of training data used for machine learning. The final function outputted is the fittest one of the population.

$$\mathrm{function}_{\mathrm{output}} = \min_{1 \le k \le M} \{\mathrm{fitness}_{\mathrm{functionTree}}, k = 1, \cdots, M\} \tag{2}$$

Where M represents the number of individuals of the entire population.

2.2 Matrix Tree

A classical entity such as the freefall of a light sphere exhibits a definite "behavior" which its trajectory can be affirmatively described by a function. However, a quantum entity such as a decaying atom doesn't exhibit any definite "behavior" which a density operator (matrix) is the only way to describe whether or not it decays probabilistically. Exactly where is the precise boundary between the classical entity and quantum entity? This is precisely the question Schrodinger posed when he came up with his cat thought experiment. By "entangling" a microscopic atom and a macroscopic cat, Schrodinger reached a paradox of a cat being dead and alive to show his skepticism about whether or not quantum mechanics is a complete theory. Schrodinger's original description [15] about his cat thought experiment is as follows:

"One can even set up quite ridiculous cases. A cat is penned up in a steel chamber, along with the following device (which must be secured against direct interference by the cat): in a Geiger counter there is a tiny bit of radioactive substance, so small, that perhaps in the course of the hour one of the atoms decays, but also, with equal probability, perhaps none; if it happens, the counter tube discharges and through a relay releases a hammer which shatters a small flask of hydrocyanic acid. If one has left this entire system to itself for an hour, one would say that the cat still lives if meanwhile no atom has decayed. The psi-function of the entire system would express this by having in it the living and dead cat (pardon the expression) mixed or smeared out in equal parts."

Essentially quantum theory is an experimental theory that fully conforms to statistical regularities (Born Rule), or simply put after repetitive observations (i.e. 100 times), approximately half the time (i.e. 51 times) the cat is alive and the other half of the time (i.e. 49 times) the cat is dead. For a single time, the experiment is performed (open the box) the cat is either dead or alive, there is no such thing as a simultaneously dead and alive cat; this can be confirmed by turning the "walls" of the sealed box transparent, the outside observer sees only a cat that is either alive or dead, and thus the final "fate" of the cat is determined by whether or not the atom decays within an hour. The randomness of whether or not the atom will decay leads to incomplete information of the state of the cat which is what causes the external observer to never be able to fully know if the cat is dead or alive. It simply cannot be inferred that from the random decay of the atom that the cat is dead and alive simultaneously, it can only be inferred that the observer is ignorant of the cat's state because of incomplete information about the decay of the atom.

To summarize there are two main points:

(1) Entity: Because the decay of the atom is undetermined, the cat is either alive or dead thus the state of the cat is uncertain.

(2) Observer: Because the cat is sealed in a non-see-through box, the observer is unable to deduce whether the cat is alive or dead due to incomplete information caused by the atom's decay or non-decay.

The uncertainty of the entity's state (the atom decays or not; the cat is alive or dead) is objective stochastic uncertainty, exactly like how a coin has the same chance of landing on either side when flipped. The stochastic uncertainty of the cats' state is just the fact that each observation produces a different outcome; while the probability of each outcome that's produced is determined by the Born rule.

Subjective epistemic uncertainty is when the observer cannot deduce the state of the cat due to incomplete information. The essentials of epistemic uncertainty (the observer guesses if the cat is alive or dead) is just the fact that the observer lacks knowledge of an event that has happened. Due to the incomplete information of the event (the observer's ignorance), the best the observer can do is to guess what may happen with a certain degree of beliefs. For example, one observer might firmly believe that the cat is 100% alive, while another might only believe that the cat is alive with 70% chance.

When observers' have to "guess" the state of cat, they are faced with having to make decisions under incomplete information constantly; furthermore, the act of scientific discovery as a whole is just a decision-making process which observers seek out the regularities of natural phenomena from the information provided by experiments. When observers are able to obtain complete information from the experimental data about a natural phenomenon, then they will be able to describe the said phenomena with a full comprehension; when observers are unable to obtain complete information from the experimental data about a natural phenomenon, then they won't quite be able to grasp the said phenomena with a complete understanding. Therefore, the most essential aspect of scientific discovery is to seek out useful information from experimental data.

Being the observer is there a way to foresee whether the cat is alive or dead without physically opening the box and peer inside to confirm?

Without looking inside the box, an approach could be the observer obtains useful information by learning historical data of whether the cat was alive or dead in previous iterations of the experiment and utilizing the obtained information to predict if the cat lives or dies; the very essence of this is just time-series forecasting under uncertainty. The challenges of time-series forecasting under uncertainty is as follows:

- Dual uncertainty: How to model both the uncertain state of the entity (cat is either alive or dead) and the epistemic uncertainty of the observer (believing if the cat is alive or dead).

- Strategy evaluation: How to find a satisfactory strategy that will allow the observer to predict the entity's state as best as possible.

2.2.1 Dual Uncertainty

By superposing all the possible states of the entity (cat is either alive or dead) and all the possible actions that the observers can take (either believe that the cat is alive or believe that the cat is dead) together according to the principle of quantum superposition [16-20] we are able to postulate an effective model of both the potential states of the entity and the collective possible actions taken by all the observers as in (3) and (4).

$$|Q\rangle = c_1|q_1\rangle + c_2|q_2\rangle \tag{3}$$

Where $|q_1\rangle$ denoting the cat is alive; $|q_2\rangle$ denoting the cat is dead. $\omega_1 = |c_1|^2$ is the objective frequency that the cat is alive; $\omega_2 = |c_2|^2$ is the objective frequency that the cat is dead.

$$|A\rangle = \mu_1|a_1\rangle + \mu_2|a_2\rangle \tag{4}$$

Where $|a_1\rangle$ denotes the observer believes that the cat is alive; $|a_2\rangle$ denotes the observer believes that the cat is dead. $p_1 = |\mu_1|^2$ are the observer's degree of beliefs that the cat lives; $p_2 = |\mu_2|^2$ are the observer's degree of beliefs that the cat dies.

The state of a complex system including a group of observers and the cat can be described as (5).

$$|\psi\rangle = c_1|q_1\rangle \otimes \prod_{i=1}^{N}|a_1^i\rangle + c_2|q_2\rangle \otimes \prod_{i=1}^{N}|a_2^i\rangle \tag{5}$$

Where N is the number of observers in the group. The density operator of the complex system can be described as (6).

$$\rho_{\text{cat+observers}} = |\psi\rangle\langle\psi| = \omega_1|q_1\rangle\langle q_1| + \omega_2|q_2\rangle\langle q_2| + \left[c_1c_2^*|q_2\rangle\langle q_1| \otimes \prod_{i=1}^{N}\langle a_1^i|a_2^i\rangle + \text{H.C.}\right] \tag{6}$$

Where the third term is a non-diagonalization term that represents the superposition of the cat being alive or dead as well as the observer being unable to deduce whether the cat is alive or dead. Observers will randomly believe that the cat is alive or dead; the observers' believing that the cat is alive or dead is "orthogonal", and when the number of observers is very large then the expectations of all the observers for whether the cat is alive or dead are then zero as (7).

$$\prod_{i=1}^{N}\langle a_1^i|a_2^i\rangle \xrightarrow{N\to\infty} 0 \tag{7}$$

(6) then becomes (8).

$$\rho_{\text{cat+observers}} \xrightarrow{N\to\infty} \omega_1|q_1\rangle\langle q_1| + \omega_2|q_2\rangle\langle q_2| \tag{8}$$

When there are a large number of observers the subjective degree of beliefs of all the observers believing that the cat is alive or dead then tends to be very close to the objective frequency of whether or not the atom decays which is determined by the Born Rule as (9).

$$\rho_{\text{cat}} = \omega_1|q_1\rangle\langle q_1| + \omega_2|q_2\rangle\langle q_2| \tag{9}$$

The cat in the box is just a classical entity, it's either alive or dead, which can be described as a mixed density operator as (9); there can't be a "superposed" cat that's alive and dead at the same time (Copenhagen Interpretation) and there isn't a cat alive in one world and dead in another (Many Worlds Interpretation).

Due to incomplete information about the decay of the atom an observer outside the box cannot fully grasp whether or not the cat is alive or dead. Since the observer won't be able know whether the cat is alive or dead without guessing, we can hypothesize that in the observer's mind the cat could be "simultaneously alive and dead" which can be described as a pure density operator in (10).

$$\rho_{\text{observer}} = |A\rangle\langle A| = p_1|a_1\rangle\langle a_1| + p_2|a_2\rangle\langle a_2| + \mu_1\mu_2^*|a_1\rangle\langle a_2| + \mu_1^*\mu_2|a_2\rangle\langle a_1| \tag{10}$$

Where p_1 are the observer's degree of beliefs that the cat is alive, p_2 are the observer's degree of beliefs that the cat is dead. The third and fourth terms in (10) are the "quantum interference" terms that indicate the observer's mind is undecided on whether the cat is indeed alive or dead, where the observer can "think" that the cat is both alive and dead.

When an observer actually "decides" on whether the cat is alive or dead, a projection of pure state to mixed state happens in the observer's mind with a certain degree of beliefs that the cat is alive or dead. The decision-making process of a single observer can be described as (11).

$$\rho_{\text{observer}} \xrightarrow{\text{Decide}} \rho'_{\text{observer}} = p_1|a_1\rangle\langle a_1| + p_2|a_2\rangle\langle a_2| \tag{11}$$

The question now becomes: is it possible to find a pure state ρ_{observer} to guide an observer to make the right choice of whether the cat is alive or dead with the highest degree of beliefs?

2.2.2 Strategy Evaluation

In this paper we develop an AI agent that simulates an observer's decision-making process which is illustrated by the projection from pure state ρ_{agent} to a mixed state ρ'_{agent}, and GP will be utilized to evolve a satisfactory pure state ρ_{agent}. The pure state is essentially just a 2x2 matrix, (11) can be described by the matrix form represented in (12).

$$\rho_{\text{agent}} = \begin{bmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{bmatrix} \xrightarrow{\text{projection}} \rho'_{\text{agent}} = \begin{bmatrix} p_1 & 0 \\ 0 & p_2 \end{bmatrix} = p_1|a_1\rangle\langle a_1| + p_2|a_2\rangle\langle a_2| \tag{12a}$$

$$|a_1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, |a_2\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix};\ |a_1\rangle\langle a_1| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, |a_2\rangle\langle a_2| = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \tag{12b}$$

Because the pure density operator ρ_{agent} is just an arbitrary 2x2 matrix, we can then approximately construct this density operator with the 8 most basic quantum gates as (13) leading it to become a "matrix tree".

$$\left\{ \begin{matrix} H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} & X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} & Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} & Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} & D = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} & T = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{bmatrix} & I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \end{matrix} \right\} \tag{13}$$

We use the addition matrix, multiplication matrix, and the logic operator or to construct a "matrix tree". The operation set and dataset of the GP algorithm for the "matrix tree" are as follows:

Operation set: $F = \{+, *, //\}$

Data set: $D = \{H, X, Y, Z, S, D, T, I\}$

After constructing an individual "matrix tree" from the operation set F and dataset D above we can then construct a population of "matrix trees", and then by using the fitness function as the evaluation criteria, the most satisfactory density matrix ρ_{agent} from the population is evolved through generations of natural selection. The "matrix tree" is essentially a decision tree that guides the AI agent which strategies to take with corresponding actions. At any given time, the expected value under the current environment (the cat is alive or dead) and the corresponding actions (the AI agent "thinks" that the cat is alive or dead) can be represented as (14).

$$\rho_{\text{cat}} = \omega_1|q_1\rangle\langle q_1| + \omega_2|q_2\rangle\langle q_2| \tag{14a}$$

$$\rho_{\text{agent}} = p_1|a_1\rangle\langle a_1| + p_2|a_2\rangle\langle a_2| \tag{14b}$$

$$\rho_{\text{cat}} \otimes \rho_{\text{agent}} = \omega_1 p_1|\langle q_1||a_1\rangle|^2 + \omega_1 p_2|\langle q_1||a_2\rangle|^2 + \omega_2 p_1|\langle q_2||a_1\rangle|^2 + \omega_2 p_2|\langle q_2||a_2\rangle|^2 \tag{14c}$$

Where (14a) is the cat observable operator; (14b) is the AI agent's observable operator; and (14c) is the composite system of the cat and the AI agent. In (14c), the first term means that the cat is alive and the AI agent "thinks" the cat is alive (this is a reward), the second term means that the cat is alive and the AI agent "thinks" that the cat is dead (this is a punishment), the third term means that the cat is dead and the AI agent "thinks" that the cat is alive (punishment), and the fourth term is that the cat is dead and the AI agent "thinks" that the cat is dead (reward). The expected value for the AI agent is the possible scenarios of what the outcome could be paired with the state of the cat that's being observed, as in (15). If the training data has N number of values, then the fitness function for the "matrix tree" is defined as (16), and it is the total sum of all the expected values of each "decision made" by the AI agent.

$$\mathrm{EV_t} = \begin{cases} \omega_1 p_1, \text{cat is alive and AI agent "thinks" so with probaility } p_1 \\ -\omega_1 p_2, \text{cat is alive and AI agent doesn't "think" so with probability } p_2 \\ -\omega_2 p_1, \text{cat is dead and AI agent doesn't "think" so with probability } p_1 \\ \omega_2 p_2, \text{cat is dead and AI agent "thinks" so with probability } p_2 \end{cases} \tag{15}$$

$$\mathrm{fitness_{matrixTree}} = \sum_{t=1}^{N} \mathrm{EV_t} \tag{16}$$

If there are M number of individuals in a population of "matrix trees", the most satisfactory “matrix tree” is the one that possess the maximum fitness function that can be described as in (17).

$$\rho_{\text{AI agent}}^{\text{output}} = \arg\max_{a}\{\mathrm{fitness_{matrixTree}}, k = 1, \cdots, M\} \tag{17}$$

By learning historical data, the more rewards that are reaped then the more accurate chance there is of predicting the next outcome of whether the cat is alive or dead. This also allows for no presumptions of the state of the cat, the more times the right outcome is “guessed” correctly by the AI agent; the best strategies and actions are effectively evolved as a result. Generation after generation of evolution, the best strategy naturally arises, which is the ultimate goal of the fitness function. The best strategy that has evolved by natural selection is the one that can be utilized to predict the future “fate” of the cat.

3 Results

The trajectory of the freefalling light sphere can be fitted by a function, while the simulated Schrodinger’s cats’ state can be probabilistically predicted by a matrix. The results presented in this section show the function and matrix produced by the AI Assistant Scientist to fit the light sphere and predict the cats’ state.

3.1 Light sphere freefall

Table 1 Data of the light sphere freefall

Label	Seconds(t)	Distance(x)
1	0.0861	1.1
2	0.1084	2.3
3	0.1250	3.4
4	0.1380	3.6
5	0.1820	6.4
6	0.2430	11.2
7	0.2460	12.8
8	0.3170	21.7
9	0.3500	24.2
10	0.3630	27.0
11	0.3810	31.4
12	0.4250	39.4
13	0.4510	42.3
14	0.4690	47.0
15	0.5200	56.8
16	0.6020	72.0
17	0.6120	73.5
18	0.7100	93.0
19	0.8120	120.1
20	0.9100	139.3
21	0.9110	141.0
22	1.0820	173.5
23	1.1810	195.0
24	1.3890	240.5
25	1.7150	317.8
26	1.9500	373.0

The data used was downloaded from the UCI Machine Learning Repository (Function Finding) as Table 1. The dataset was divided into two parts: training dataset (1-20 datapoints) and verify dataset (21-26 datapoints). The fitting results of the training dataset is shown in Figure 1. The blue line is the original data, yellow line is the calculated data. The simulated distance of the freefalling light sphere is calculated by the "function tree" in (18).

$$x = e^{6\sin\sqrt{t}} \tag{18}$$

The predicted distance of the freefalling light sphere is calculated by the "function tree" in (18), shown in Figure 2. The blue line is the original distance, the yellow line is the predicted distance. Table 2 shows the observed data and calculated data by the "function tree"; the Mean Absolute Percentage Error (MAPE) of the calculated distance is 4%.

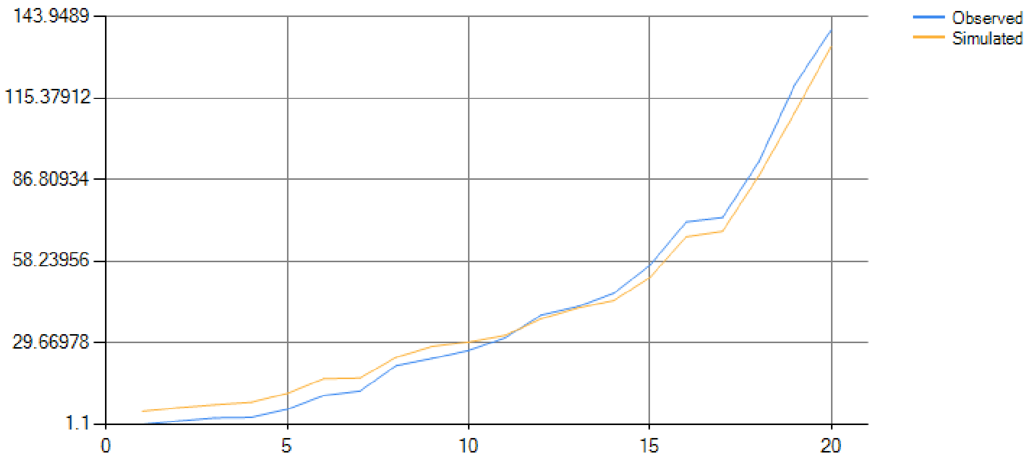

Fig. 1 The fitting results of the training dataset. The horizontal axis represents the label of each datapoint.

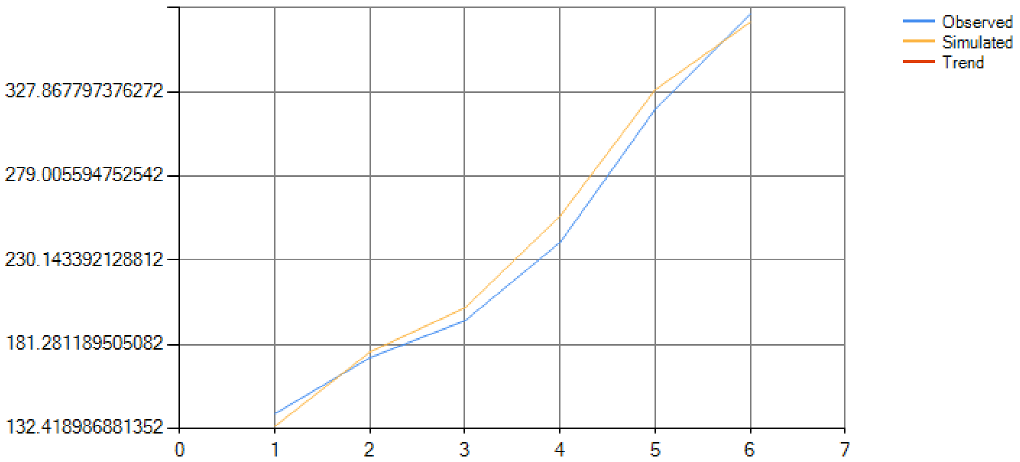

Fig. 2 The verify results of the freefalling light sphere.

Table 2 Verify results of the light sphere freefall

Label	Calc. Dis.	Org. Dis.
21	133.756	141.0
22	176.798	173.5
23	202.488	195.0
24	255.784	240.5
25	329.132	317.8
26	368.342	373.0

3.2 Schrodinger's Cat

The outside observer cannot decide the fate of the cat inside the box, it is whether or not the atom decays that decides if the cat lives or dies. From inside the box the cat is either alive or dead, but because whether or not the atom decays causes incomplete information, and since the box is sealed an outside observer can’t fully grasp the state of the cat until the observer opens the box and looks inside to see whether the cat is actually alive or dead. If there is N number of boxes all set up containing a cat in accordance to Schrodinger's thought experiment (Geiger counter, radioactive substance, hydrocyanic acid), an outside observer would have to open every single one individually to see if the observer's original "guess" regarding the state of the cats' in each box is right or wrong, thus in turn the observer's judgement about whether each cat in every box is alive or dead then becomes a decision-making under uncertainty scenario. In this paper, we generate 25 boxes of "Schrodinger's cats" each labeled from 1 to 25, then we use a "digital coin" to simulate the decay or non-decay of the atom, when it "lands" on heads that means the atom decays (cat dies) while if it "lands" on tails then the atom does not decay (cat lives). The time series of simulated Schrodinger's cat thought experiments labeled 1 to 25 are shown in Table 3, 0 being that the cat is observed as alive and 1 that the cat is observed as dead.

Table 3 Data of simulated Schrodinger’s Cat thought experiment

Label (N)	Cats’ state (q)
1	0
2	1
3	0
4	1
5	0
6	0
7	0
8	1
9	0
10	0
11	0
12	1
13	0
14	1
15	1
16	0
17	1
18	1
19	1
20	0
21	1
22	1
23	0
24	1
25	1

The dataset was divided into two parts: training dataset (1-19 datapoints) and verify dataset (20-25 datapoints). We ran two consecutive training sessions on the training dataset. In both sessions, the training parameters were set the same; with 3 AI agents and a training frequency of 1000 repetitions. An optimized "matrix tree" is obtained after each session; this “matrix tree” is then utilized to produce 6 predictions whether the cat is alive or dead on the verify dataset, in which the final prediction regarding the fate of the cat is the result of 3 AI agents cooperating as in Table 4. For the specific forecast values in Table 4, please refer to the supplementary materials.

Table 4 Prediction results of session 1 and session 2

		Session 1						Session 2					
Label	Cats’ state	1	2	3	4	5	6	1	2	3	4	5	6
20	Alive	1	1	1	1	0	0	0	1	1	1	1	0
21	Dead	1	1	1	0	1	0	1	1	0	1	0	0
22	Dead	1	0	1	0	1	1	0	1	1	1	1	0
23	Alive	1	0	0	0	0	0	0	1	0	1	1	1
24	Dead	0	1	0	1	0	1	0	1	1	1	0	1
25	Dead	1	1	0	1	0	0	1	0	1	0	1	1

In Table 4, 0 represents that the AI agents “believe” that the cat is still alive, and 1 represents that the AI agents "believe" that the cat has died. For example, {1, 1, 1, 1, 0, 1} from the first possible forecast of session one represents the action sequence predicted through the cooperation of three AI agents that "believe" the state of the cat will be {Dead, Dead, Dead, Dead, Alive, Dead}, the odds of this forecast are 50%, and the rest of the 11 possible forecasts can be deduced in the same way. The average odds of all 12 forecasts are 54.166% which is very close to the 50-50 frequency of the decaying of the atom (Born Rule). By empirical means, we've found that when the accuracy of the 12 single forecasts is more than 50% on average, then the forecast of the final action sequence obtained by means of majority rules will be greatly more accurate. For example; for the state of the cat labeled 20 in Table 4, the forecast outcomes of the 12 individual action sequences are {1,1,1,1,0,0,0,1,1,1,1,0}, in which of the 12 individual forecast outcomes only 4 "believe" that the cat is still alive, while 8 "believe" that the cat is already dead, thus according to majority rules the final forecast is "believe" that the cat is dead. For the cats labeled 21-25, the same rules can be applied to deduce a final action sequence for forecasting the state of the cat, as in Table 5.

Table 5 Final forecast outcomes by majority rules

Label	Observed cats' state	Predicted cats' state
20	0	1
21	1	1
22	1	1
23	0	0
24	1	1
25	1	1

In Table 5, 0 represents that the cat is alive, 1 represents that the cat is dead, and the odds of the final action sequence (obtained by applying majority rules) utilized to forecast the state of the cat are 83%. Therefore, the 83% odds illustrate that even if an external observer doesn't open the box and physically look inside to see if the cat is alive or dead, through the cooperation of 3 AI agents, the observer can somewhat "guess" correctly the state of the cat. Of course, even though the odds for the 6 short-term forecasts of the cats labeled 20-25 was 83%, but for a longer forecast horizon the accuracy rate will still be closer to 50-50. Our strategy when it comes to forecasting is to study more recent historical data, i.e. the cats labeled 1-19, to then make predictions about the next 6 outcomes, i.e. cats 20-25. We believe that the outlying values of the past and the to-happen values of the distant future for short-term forecasts are not of much significance as they are more useful for statistical analysis and other mathematical means; our strategy is to learn valuable information of the recent past to then forecast the near foreseeable future.

4 Discussion

For the freefalling light sphere, the quantum-like evolutionary algorithm is able to produce a near perfect (96% accurate) function to predict the light sphere's past and future trajectory because definite information of this natural phenomenon can be obtained from the observed data. Regarding Schrodinger's Cat, due to the lack of enough definite information (the atom randomly decays yet we still don't know why mother nature does so), the quantum-like algorithm cannot produce a definite function that can accurately state whether the cat is alive or dead, but instead is able to produce a matrix (projection operator) which reconstructs the historical states of the cats' in the prior boxes (labeled 1-19) and then gives a satisfactory probabilistic prediction of what the cats' state (alive or dead) are in the next few boxes that are opened (labeled 20-25).

Regarding Schrodinger's Cat specifically, and quantum theory in general, there has been an ongoing debate for more than 100 years, and this debate will very likely continue [21-36]. Up until now there has been three major interpretations of the Schrodinger's cat paradox:

(1) Copenhagen Interpretation: There is no point in discussing whether the cat is dead or alive without opening the box and looking (measuring); according to mathematical models the cat inside the box is in a superposed state and it is only when an external observer opens the box to look inside that the state of the cat "collapses" from a superposed state to a

state of being only alive or dead.

(2) Many-worlds Interpretation: A physical macroscopic cat cannot be in a superposed state of dead and alive, thus there is no so-called "collapse" of what state the cat is in, but the act of looking (measurement) at the cat will cause a split into two parallel worlds; one in which the cat is alive and one in which the cat is dead.

(3) QBism Interpretation: It is practically meaningless to discuss whether the cat in the box is alive or dead without an external observer opening it and looking inside (measure), what's more meaningful is what expectation the observer has in their mind when they open the box and looks to see if the cat inside is alive or dead, and based on the new information an observer obtains regarding the state of the cat they then update their beliefs about whether the cat indeed lives or dies.

The big conundrum of the Copenhagen Interpretation is the "wave packet collapse" (the quantum measurement problem); The issue with the many-worlds interpretation is the "splitting" of the many parallel worlds, but the bigger issue is that the many-worlds interpretation assumes the entire universe in existence is a huge complex wave function, and it is because of this assumption the Archilles heel is that no one can either prove this to be correct but at the same time it cannot be proven false, thus is a question that is beyond the very scope of the modern-day scientific methodology and current research; In a sense the Quantum Bayesian (QBism) interpretation somewhat avoids the "wave packet collapse" (the problem that the Copenhagen Interpretation faces), and avoids the "splitting of parallel worlds" (the challenge of the many-worlds interpretation) as well, where QBism believes the key to solving the Schrodinger's Cat paradox is being how much and what information the observer can grasp about the state of the cat, however the challenge is then how can QBism practically calculate the observer's degree of expected beliefs and then accurately predict whether the cat is alive or dead.

In our proposed interpretation of Schrodinger's Cat based on a quantum-like decision theory, we subtly fused together the advantages and somewhat avoided the shortcomings of the three interpretations mentioned above. The cat in the box of Schrodinger's thought experiment is just like any normal "classical" cat, inside it's only either alive or dead. For the observer as well, the external observer is also just a "classical" average Joe, someone who "measures" the state of the cat by opening the box and records information on whether the cat is alive or dead; however the act of opening the box by the observer in itself does not cause the cat to "collapse" from a superposed state of alive and dead to one or the other state, neither does it cause the cat to "split" into different parallel worlds where it lives in one and dies in the other. Of course, the external observer as a scientist who has not opened the box to see whether the cat inside is alive or dead yet will come up with various theories to "guess" the exact state of the cat in the box, that is by testing their hypothesis' through repeated experimentation with the aim of being able to specifically reconstruct historical experimental data and to make predictions about the future. In this paper the quantum-like evolutionary algorithm we proposed obtains a satisfactory "theory" by machine learning historical experimental data through the cooperation of multiple AI agents' resulting in the successful reconstruction of the historical data on the cats' past states and sufficient predictions about the cat's future states.

Regarding quantum entities we shouldn't persistently ask "Does God play dice with the universe? (Determinism)" and "Does the moon exist when nobody looks at it? (Realism)", but instead we should be asking how much useful information can be obtained from the interactions between the measuring instruments and the quantum entities, and more importantly how to use the information obtained to reconstruct the historical states of the quantum entities and how to predict their future states in the best way possible. This paper has attempted to study this problem from an information perspective, by using a quantum-like decision theory to reconstruct historical events and to predict future events. As mortal human observers resident within the vast universe, we cannot propel ourselves to leap out of the universe to get a glimpse of the universe as a whole to formulate a theory of everything (obtain complete information of the entire universe); the best we can do is with the help of AI assistants to "play a game" with nature (or maybe even attempt to play dice with God, if God really does indeed play dice with the universe) by evolving a satisfactory enough theory through means of Darwinian natural selection to grasp as much useful information about the universe as possible on our quest to understanding the extragalactic cosmos.

Declarations

Data availability. The freefall light sphere data used is part of the dataset Function Finding that is publicly available for download from the University of California Irvine (UCI) Machine Learning Repository. The data representing the decay or non-decay of the atom for the simulated Schrodinger's Cat thought experiment was generated by the authors. The authors confirm that all data that support the findings of this study are included within the main text manuscript and its supplementary materials document. All data are also available from the corresponding author K.X., upon reasonable request.

Authors' contributions. All authors conducted the research and contributed to the development of the model. L.X. contributed to the research from the aspects of machine learning, decision theory, and quantum theory. K.X. prepared the data, did data analysis, and wrote the main manuscript. All authors reviewed the manuscript.

Conflict of interests. The authors are affiliated with XINVISIONQ, INC., the developer of a commercial forecasting tool, but for transparency of the experiments conducted in this study, all data generated and produced are included in the paper's main text and supplementary materials document. The results and conclusions presented in this paper are in accordance with the ethical requirements of academic research as no data was manipulated or manually adjusted. The authors have no other competing academic interests to declare, as the methodology formulated and findings reached in this paper are the authors' original research work.

Funding. This work received no funding.

References

[1] Guimerà, Roger, Ignasi Reichardt, Antoni Aguilar-Mogas, Francesco Alessandro Massucci, Manuel Miranda, Jordi Pallarés and Marta Sales-Pardo. "A Bayesian machine scientist to aid in the solution of challenging scientific problems." Science Advances 6 (2020).

[2] Langley, Pat. "Rediscovering Physics with BACON.3." International Joint Conference on Artificial Intelligence (1979).

[3] Schmidt, Michael D. and Hod Lipson. "Distilling Free-Form Natural Laws from Experimental Data." Science 324 (2009): 81 – 85.

[4] Brunton, Steven L., Joshua L. Proctor and J. Nathan Kutz. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems." Proceedings of the National Academy of Sciences 113 (2015): 3932 – 3937.

[5] Lemos, Pablo, Niall Jeffrey, M. Cranmer, Shi-Lin Ho and Peter W. Battaglia. "Rediscovering orbital mechanics with machine learning." ArXiv abs/2202.02306 (2022).

[6] Udrescu, Silviu-Marian and Max Tegmark. "AI Feynman: A physics-inspired method for symbolic regression." Science Advances 6 (2019).

[7] Cornelio, C., Dash, S., Austel, V. et al. Combining data and theory for derivable scientific discovery with AI-Descartes. Nat Commun 14, 1777 (2023). https://doi.org/10.1038/s41467-023-37236-y.

[8] Krenn, M., Pollice, R., Guo, S.Y. et al. On scientific understanding with artificial intelligence. Nat Rev Phys 4, 761–769 (2022). https://doi.org/10.1038/s42254-022-00518-3.

[9] Xin, L., Xin, H. Decision-making under uncertainty – a quantum value operator approach. Int J Theor Phys 62, 48, https://doi.org/10.1007/s10773-023-05308-w (2023).

[10] Xin, L.; Xin, K.; Xin, H. 2023 On Laws of Thought—A Quantum-like Machine Learning Approach. Entropy, 25, 1213. (doi:10.3390/e25081213)

[11] Holland, J., Adaptation in Natural and Artificial System, Ann Arbor, MI: University of Michigan Press (1975).

[12] Goldberg, D.E., Genetic algorithms – in search, optimization and machine learning, New York, NY: Addison-Wesley Publishing Company, Inc. (1989).

[13] Koza, J.R., Genetic programming, on the programming of computers by means of natural selection, Cambridge, MA: MIT Press (1992).

[14] Koza, J.R., Genetic programming II, automatic discovery of reusable programs, Cambridge, MA: MIT Press (1994).

[15] Trimmer, John D. "The Present Situation in Quantum Mechanics: A Translation of Schrödinger's 'Cat Paradox' Paper." Proceedings of the American Philosophical Society 124, no. 5 (1980): 323–38.

[16] Von Neumann, J. Mathematical Foundations of Quantum Theory, Princeton, NJ: Princeton University Press (1932).

[17] Dirac, P.A.M., the Principles of Quantum Mechanics, Oxford University Press (1958).

[18] Heisenberg, W., the Physical Principles of the Quantum Theory, Chicago, IL: The University of Chicago Press (1930).

[19] N. S. Yanofsky, M. A. Mannucci, Quantum Computing for Computer Scientists, Cambridge University Press, Cambridge, 2008.

[20] Nielsen, M.A. and Chuang, I.L., Quantum computation and quantum information, Cambridge University Press (2000).

[21] Wheeler, J. A. and Zurek W.H., Quantum Theory and Measurement, Princeton, NJ: Princeton University Press (1983).

[22] Jammer, M., the Philosophy of Quantum Mechanics, New York, NY: Wiley (1974).

[23] Rovelli, C., Helgoland: Making Sense of the Quantum Revolution, New York, NY: Riverhead Books (2021).

[24] Zurek W.H., Decoherence, einselection, and the quantum origins of the classical, Rev. Mod. Phys. 75, 715-75 (2003).

[25] Popper, K.R., Quantum Mechanics without "the Observer", In Bunge, M. (Ed.). Quantum Theory and Reality, Springer (1967).

[26] van Fraassen, B. C., Hidden Variables and the Modal Interpretation of Quantum Statistics, Synthese 42, 155-65 (1979).

[27] Everett, H., "Relative State" Formulation of Quantum Mechanics, Rev. Mod. Phys. 29, 454 (1957).

[28] Fuchs, C.A, Schack, R., Quantum- Bayesian coherence, Rev. Mod. Phys. 85, 1693 (2013).

[29] Mittelstaedt, P., the Interpretation of Quantum Mechanics and the Measurement Process, Cambridge: Cambridge University Press (1998).

[30] Omnes, R., the Interpretation of Quantum Mechanics, Princeton, NJ: Princeton University Press (1994).

[31] Bub, Jeffrey. Interpreting the Quantum World, Cambridge: Cambridge University Press (1997).

[32] Baggott, Jim., Beyond measure: Modern physics, philosophy, and the meaning of quantum theory, Oxford: Oxford University Press (2004).

[33] Healey, R. A., and Hellman G., Quantum Measurement: Beyond Paradox, Minneapolis, MN: University of Minnesota Press (1998).

[34] Bunge, M., Quantum Theory and Reality, Springer (1967).

[35] Maudlin, T., Philosophy of Physics: Quantum Theory, Princeton, NJ: Princeton University Press (2019).

[36] Albert, D.Z., Quantum Mechanics and Experience, Cambridge, MA: Harvard University Press (1992).