Parabola Volume 59, Issue 1 (2023)

Introduction to zero-knowledge proofs

Trevor Chi-Yuen Tao[1]

1 Introduction

In this report, I explain the concept of zero-knowledge proofs and offer an interesting application for detecting the use of engine-assistance in competitive chess.

2 Zero-knowledge proofs

A zero-knowledge proof describes a situation where a one person can convince another that some statement is true, but without giving away any further information. Most proofs require giving away some information. For instance, if Peggy[2] claims that a Sudoku puzzle has multiple solutions, then she can prove this to Victor by writing down (at least) two solutions. But that is not a zero-knowledge proof since Victor now possesses knowledge of these solutions, which he didn't have previously.

Zero-knowledge proofs are typically achieved via interaction between two people in the form of challenges. If Peggy claims that some fact is true and she can solve a sufficient number of challenges, then Victor will conclude Peggy's claim is correct.

A simple and well-known example of a zero-knowledge proof concerns a cave and a magic door that requires knowledge of a password to open; see Figure 1. Peggy randomly chooses between two paths A and B. Victor then names one of the two paths at random and Peggy must return from the named path. Peggy can succeed with certainty if she knows the password, or with probability 0.5 if she doesn't. If Peggy succeeds a large number of times, then she can convince Victor she indeed knows the password [1].

A number of more serious applications exist. One such application concerns cycles in graph theory. Given a graph G with k nodes, a cycle is a path $n_0, n_1, \ldots, n_k$ where every pair of adjacent nodes is connected by an edge, where $n_0 = n_k$ and where every other node appears exactly once. Such a cycle is called a *Hamiltonian* cycle, and proving the existence or otherwise of such a cycle is known to be NP-complete [4].

The zero-knowledge protocol for that application works as follows: Peggy creates a graph H that is different to G. Victor then randomly chooses one of two challenges:

[1] Trevor Chi-Yuen Tao has a PhD in applied mathematics. He is a research scientist currently working for the Australian Department of Defence. Trevor is a keen chess and scrabble player and his other hobbies include mathematics and music. Trevor is the brother of world-renowned mathematician Terence Tao.

[2] When discussing zero-knowledge proofs, the names Peggy and Victor are typically used since P and V stand for Prover and Verifier, respectively.

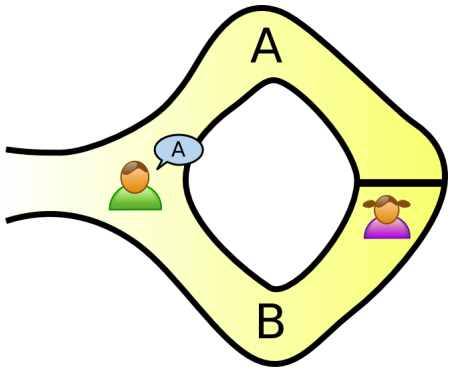

Figure 1: Cave and magic door.
Image source: https://commons.wikimedia.org/w/index.php?curid=313645.

either show a cycle in H or give a 1-1 correspondence between nodes in G and H to prove that they are isomorphic. This challenge is similar to the cave example because Peggy succeeds with certainty if she knows a Hamiltonian cycle; otherwise, she succeeds only half the time [2].

3 Detecting engine-assistance in chess

3.1 Expert vs engine-assisted novice

Suppose that Victor is a chess novice and his friend Peggy is an expert. Victor has access to a very strong chess engine; that is, a computer equipped with chess-playing software. Peggy claims that she can reliably distinguish between the play of a novice versus the play of a strong engine. To test this theory, Victor and Peggy agree to play a 20-game match online with the following conditions:

(a) Normal chess rules apply.

(b) In each round, Victor must commit to one of the following strategies:
(1) playing to the best of his ability without engine assistance;
(2) at every turn choose among the top three moves given by the engine.

(c) If Peggy's King occupies one of the four centre squares (d4,d5,e4,e5), then she has the option of aborting the game and calling Victor. If Victor was using an engine, then Peggy wins; otherwise, Victor wins. An example game state with White king on e4 is shown in Figure 2.

In other words, a game of chess can end normally in several ways (checkmate, stalemate, resignation, etc.) or if Peggy occupies the centre with her King and elects to call Victor. Note that nothing special happens if it is Victor's king who occupies the centre.

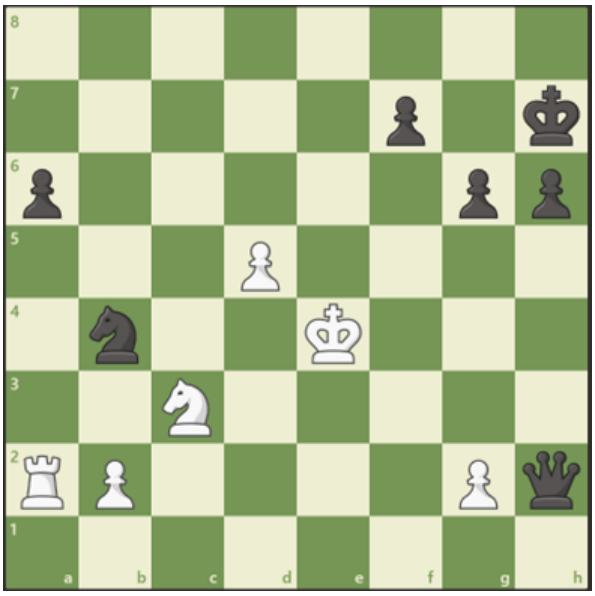

Figure 2: Game state with White King on e4.

In this report, I will use the term "King of the Hill" to describe the situation where a king occupies one of the four centre squares. This is based on a well-known chess variant of the same name [3].

I now discuss four possible scenarios based on whether Peggy believes that Victor is engine-assisted and whether Victor is actually engine-assisted. I use the notation **X-Y** to mean "Peggy believes X and the truth is Y", where "Human" means Victor is playing unassisted and "Engine" means Victor is using an engine.

Human-Human: Peggy wins on merit.

Human-Engine: Peggy loses to the engine.

Engine-Human: Victor allows Peggy achieve King of the Hill. Peggy calls and loses.

Engine-Engine: Victor allows Peggy achieve King of the Hill. Peggy calls and wins.

These scenarios are summarised in the table below.

Belief \Actual	Human	Engine
Human	Peggy wins on merit	Peggy loses to engine
Engine	Peggy calls Victor and loses	Peggy calls Victor and wins

3.2 Assumptions

Before proceeding, I wish to discuss a number of assumptions:

(a) I will assume that if Peggy plays to the best of her ability, then it will be difficult (but not impossible) to achieve King of the Hill.

(b) I have also ignored the possibility of draws. I also assume the difference in playing strength is such that Peggy always beats Victor but always loses to the engine.

(c) I will assume Victor only has a "binary strategy" of using engine assistance for every move or not at all. In practice, a savvy cheater may choose to only use engine assistance for some but not all of his moves, to make detection more difficult.

(d) The King of the Hill variant is clearly biased if the win condition only applies to one player. I am assuming that this bias outweighs the opponent's ability to use an engine. I also assume the engine does not recognise the concept of King of the Hill.

The first assumption requires some explanation: centralising the king is generally not relevant to the goal of achieving checkmate.[3] It is easy enough to construct individual positions where centralising the King happens to be a correct strategy, but this would never hold up in the long run, e.g., in a 20-game match. The important point is that Peggy cannot "have it both ways". Suppose that Peggy plays to the best of her ability until the last minute, when she recognises that defeat is inevitable. When the last minute occurs, it will usually be impossible for Peggy to change strategy and hope to save herself with King of the Hill. Therefore, Peggy is incentivised to call Victor for cheating only when she has credible evidence of cheating - not because she has nothing to lose!

3.3 The chess speaks for itself

If the above assumptions hold, then the 20-game match described above can be thought of as a protocol for zero-knowledge proofs. If Peggy wins all of her games against Victor, then it is reasonable for Victor to conclude that Peggy's claim is correct. If Peggy loses even one game, then her claim is refuted. Note also that Peggy does not have to explain why she thinks Victor is playing with or without engine assistance. All Peggy has to do is (dare I say it) to let the chess speak for itself.

[3] In contrast, a task such as "capturing at least six enemy pieces" is relevant, since obtaining a superiority in material is a fundamental strategy that is obvious even to weak players.

4 Competitive Chess in Practice

The previous section described a sanitised thought experiment. One attraction of zero knowledge proofs is that many strong players can "sense" when a player is cheating, but they have difficulty explaining to the layman the basis of their suspicions. However, a number of the above assumptions will not hold up in practice. Moreover, most tournaments involve a large number of players and there is no prospect of playing the same opponent more than once. Nevertheless, I believe that an interesting rule can be used in practice to deter the use of engine assistance in chess. Unfortunately, I am unable to support this idea with an actual experiment. The main point of this article is to illustrate the concept of zero-knowledge proofs.

My idea is the following new rule:

(a) If a player achieves King of the Hill during a game, then he has the option of demanding that both players be tested for electronic doping without fear of any penalty for a false accusation, such as loss of the game, revoking of titles etc. Obviously, King of the Hill applies to both players, unlike the thought experiment described above.

(b) If a player cannot achieve King of the Hill but still wishes to call his opponent, then he must be willing to risk severe penalties for an incorrect accusation.

(c) There is no obligation to call an opponent if the player does achieve King of the Hill.

As an aside, some chess players may recognise Figure 2 from the famous game [4] between Indian Grandmaster Viswanathan Anand and Nikhil Kamath, a co-founder of Zerodha, a stock-broking company. Nikhil was forced to admit to cheating after Anand resigned. Actually, I have taken the liberty of adding a few extra moves: instead of resigning, White marches his King towards the centre and makes no effort to win the game by normal means.

[4] https://www.chessbase.in/news/Vishy-Anand-checkmate-covid-simul-game-against-Nikhil-Kamath

References

[1] J.-J. Quisquater, M. Quisquater, M. Quisquater, M. Quisquater, L. Guillou, M.A. Guillou, G. Guillou, A. Guillou, G. Guillou, S. Guillou and T. Berson, How to explain zero-knowledge protocols to your children, *Lecture Notes in Computer Science* **435** (1989), 628–631.

[2] M. Blum, How to prove a theorem so no one else can claim it, in *Proceedings of the International Congress of Mathematicians, Berkeley, California, USA, 1986*, pp. 1444–1451, 1987.

[3] M.A. Gehrke, *Assessing Popular Chess Variants Using Deep Reinforcement Learning*, PhD thesis, Podunk University Arcana Department, 1996.

[4] M. Hosseininia and F. Dadgostari, Hamiltonian paths and cycles, in R.Z. Farahani and E. Miandoabchi (eds.), *Graph Theory for Operations Research and Management: Applications in Industrial Engineering*, pp. 96–105, IGI Global, 2012.