TANS: A Chess-Inspired Notation System for Strategy Analysis of Tennis Games

Yuexi Song[1], Chuanfei Li[2], Hao Cao[3], Ling Wu[4], Huanhuan Zheng[1], and Zhenkai Liang[1]

[1] National University of Singapore
[2] Raffles Institution
[3] Nanyang Technological University
[4] ChemT Biotechnology

Abstract. Tennis is a challenging sport that requires a comprehensive set of skills, including physical abilities, technical proficiency, and mental resilience. More importantly, tennis is also a game of strategy, demanding players to constantly evaluate the situation, calculate the shot placement and tactics, and control the flow of match. Chess, another game of strategy, has long benefited from the standard notation system, which enables deep and systematic analysis of game strategy in chess and even AI integration. In contrast, tennis currently lacks a structured and standardized framework for analyzing strategic game play. To bridge this gap, this paper presents Tennis Algebraic Notation System (TANS), a novel tennis notation system inspired by traditional chess notations. Unlike existing notation methods, TANS makes it easy for automated data extraction and analysis of game strategy and winning patterns. Using charted data with the new notational system, we can carry out analysis of match statistics and winning strategy analysis of tennis doubles matches. By summarizing tennis matches in a concise and information-rich notation system, we believe that TANS represents an important step toward enabling future systematic analysis of tennis strategy and AI-based game simulation.

Keywords: Strategy Analysis, Notation System, Tennis Analytics

1 Introduction

Tennis is a challenging sport that requires a comprehensive set of skills, including physical abilities, technical proficiency, and mental resilience. More importantly, tennis is also a game of strategy, demanding players to constantly evaluate the situation, calculate the shot placement and tactics, and control the flow of match. Data analytics is widely used in tennis games to give overall view of the game and players. However, there is only a limited amount of research in automated strategy analysis in tennis [14].

Chess is another game of strategy. Chess has a mature notation system for game charting and analysis [5], which also enables extensive analysis and modeling [18], including AI-based chess players. In contrast, a key missing piece in tennis is a notation system that precisely and concisely describes tennis moves. The challenge is that tennis movements are much more complex than those in chess. Tennis games are influenced by ball speed, spin, player positions, etc. It is not straightforward to include these features into a concise description.

However, we observed that *chess and tennis have many similarities in game strategy*. For example, both games have the notion of controlling space, planning moves, etc. Motivated by this similarity, we view tennis as a board game in which *one player moves other players through the placement of tennis balls*. As the tennis ball is the cause of player movements, we use the ball's position as the primary focus in our notation, which is comparable to the piece positions of the chess game.

In this paper, we define TANS, Tennis Algebraic Notation System. Inspired by the algebraic notation system [5] of chess, TANS describes detailed moves in tennis games in a concise format. TANS is designed to support both tennis analytics and strategy analysis. Based on TANS, we adopted algorithms, such as Grey Relational Analysis (GRA), to identify game strategies, showcasing the capability of TANS in enabling fine-grained strategy analysis.

In our evaluation, we focus on tennis doubles games when carrying out strategy analysis, as tennis doubles games have stronger strategy patterns. Using the data charted from matches such as Australian Open Doubles Final, we showed that TANS can support standard tennis analytics tasks. In addition, it can reveal deeper insights in winning patterns. We analyzed the effectiveness of employing the "I-formation" in tennis doubles and identified winning patterns for the serving team, which shows the strongest correlation with winning the point.

In summary, TANS records tennis matches in a concise and precise notation system. We believe that TANS represents an important step toward enabling future systematic analysis of tennis strategy and AI-based game simulation. In particular, we made the following contributions in this paper.

- We define TANS, a novel tennis notation system. Inspired by chess notation systems, we define the syntax of the notation language to record tennis moves.
- We develop applications on top of TANS to show that TANS can support deeper tennis analytics and strategy discovery.
- We evaluated the TANS language and applications using past professional games, where we discovered interesting strategy patterns from charted games.

2 Background and Related Work

2.1 Tennis Strategy Analysis

Tennis is a highly strategic sport, and over the years, researchers have made continuous efforts to model game play and analyze winning strategies. As a start, Chiu et al. [8] developed a mathematical model and utilized math and physics simulation in identifying optimal strategy for tennis. Vis et al. [22] used descriptive variables such as stroke type, player position and used data mining for pattern discovery in tennis rallies. Martínez-Gallego et al. [16] utilized some other descriptive variables such as result of the last shot of the game, shots per game, and analyzed the game structure and point ending characteristics of tennis doubles matches. Similarly, Torres-Luque et al. [23] designed an observational instrument for obtaining objective information about singles matches for analysis. Kocib et al. [12] observed 8 sets in 18 professional men's doubles matches to examine the frequency and effectiveness of various tactical formations. Focused on tennis double games, Liu et al. [15] developed a Markov-based framework for modeling doubles game play. Given the additional complexities of doubles tennis, including player coordination and dynamic positioning, this work represents an important step toward enhancing the strategic understanding of doubles matches. These studies provided valuable insights into tennis strategy analysis. However, different research works have come up with different formulations of tennis games, requiring manual annotation of tennis games for each research work. Thus a major limitation to these studies has been the lack of comprehensive match charting data.

With the advent of the Match Charting Project, this limitation has been significantly mitigated. The project, hosted on TennisAbstract.com [7], provides thousands of data points from professional matches, charted based on their predefined charting language. Carlo, et al. [19] utilized the language and developed an AI Framework for tennis singles games. In contrast to mainly statistically analysis work done previously, their work opened a new door for utilizing AI simulation in identifying game winning strategies. However, the charting language from TennisAbstract was originally designed for ease of recording rather than for machine learning application, limiting the accuracy of the framework. Hence, a key missing component across these works is a systematic notational language specifically designed to support strategy analysis of tennis games.

2.2 Chess Notation Systems

There have been significant developments in analyzing chess game winning moves and patterns. Chess has a well structured notation system which allows powerful computational tools to analyze chess games and dive deep into strategic exploration. A groundbreaking development in this area is AlphaZero [21], which is a chess AI that mastered the game through reinforcement learning in self-play. Later, Ruoss et al. [20] optimized performance through reinforcement learning to generate moves without relying on explicit search. A basis for these advancements is the algebraic notation [5], the official notation system of chess. It includes the following components.

1. **Chessboard Coordinates.** The board has files (columns) labeled a–h (left to right). The ranks (rows) are numbered 1–8 (bottom to top), shown in Figure 1.
2. **Piece Abbreviations.**
   - `K`: King
   - `Q`: Queen
   - `R`: Rook
   - `B`: Bishop
   - `N`: Knight
   - No letter for pawns (just the destination square)
3. **Move Notation.**
   Write the piece abbreviation followed by the destination square. For example, `e4` (pawn to e4); `Nf3` (knight to f3); `Bb5` (bishop to b5).
   Captures. Use `x` between the piece and the destination. For example, `Nxe5` (knight captures on e5); `Rxd7` (rook captures on d7); `exd5` (pawn from File e captures on d5).

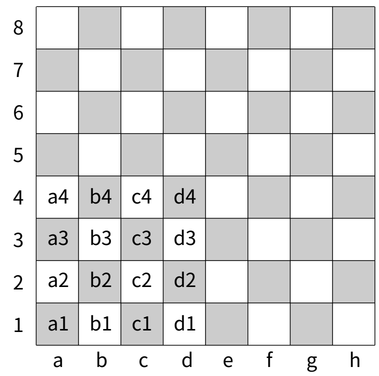

Fig. 1: Chessboard and coordinates.

Special Moves. *Castling*: Kingside: `O-O`, Queenside: `O-O-O`. *Pawn promotion*: `e8=Q` (pawn reaches e8 and promotes to a queen). *En passant*: `exd6 e.p.`
Extra notations If there are two of the same piece that can move to the same square, specify which one moved by adding: The file it came from (`Rea8` means the rook from the "e" file moved to a8). If both pieces are on the same file, use the rank instead (`R1d3` if rooks are on d1 and d5, and the one from d1 moves to the d3 square).
Check and Checkmate. *Check*: Add + behind the piece abbreviation and destination square, e.g., `Qg5+`.
Checkmate: Add # behind the piece abbreviation and destination square, e.g., `Qh8#`.

4. **Game End Notations.**
   `1-0` (White wins).
   `0-1` (Black wins).
   ½–½ (Draw).

The algebraic notation is the standard representation to describe moves. Additional information, such as player names, dates, are included as tags at the beginning of the description to form the Portable Game Notation (PGN) [9] for chess. PGN is widely used by chess software and chess game databases.

3 Tennis Algebraic Notation System

Inspired by chess notation systems, we developed a tennis algebraic notation system (TANS) to enable systematic analysis of tennis strategies.

The detailed description of the language is as follows:

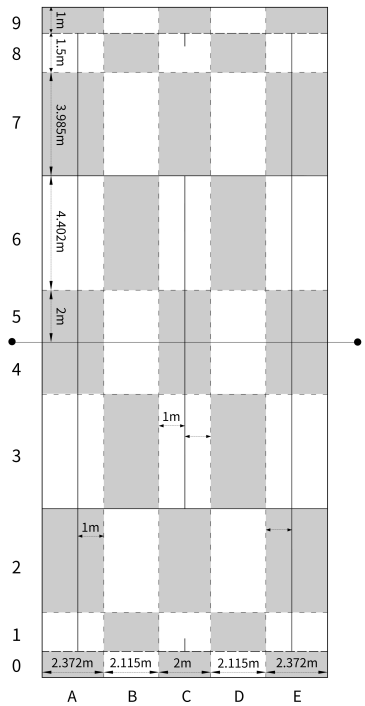

Fig. 2: Tennis court coordinate in TANS.

1. **Court Coordinates.**
   A naive method to assign court coordinate is to equally divide the court, similar to the chess board. However, unlike in a chess board, the regions in the tennis court are of different importance in tennis strategy. Existing studies such as [17] have shown that ball landings occur more frequently in certain areas of the court, and winning shots are often associated with specific regions, such as near the net or close to the sidelines. Therefore, we made a coordinate system based on the *strategic meaning* of tennis courts.

   TANS Coordinates. As illustrated in Figure 2, the solid lines are the court partitions where the tennis court is divided into a grid of ten rows (ranks) 0-9 by five columns (files) A-E, forming 50 distinct sections. Each section is uniquely labeled with letters and numbers, such as “C4”. In particular, Row 0 and Row 9 are reserved for player positioning behind the baseline where players prepare to serve, receive, or respond to long shots.

This coarse-grained court partition is designed for easier charting by humans. It captures the overall position of the ball on a court, such as left/right or center. In case a finer grained partition is needed, such as for computerized processing, the court columns (files) can be further divided, which we illustrate in the Appendix B.

2. **Player Abbreviations.**

   To identify shots made by each player, tags are used to label the players on court. Players are labeled as X and Y in singles games and as X, Y, Z, and W in doubles games, where X and Y belong to one team, and Z and W belong to the opposing team. The players' names can be included in the metadata section at the beginning of the charting data, similar to the Portable Game Notation (PGN) file format used in chess.

3. **Shot and Move Notation.**

   We view the tennis game as players moving the tennis ball strategically. Hence, each shot is encoded in the following format: `[Player tag] + [Stroke] + [Ball landing position]`. For example, `WfB6` stands for Player W hits a forehand ground stroke and the ball lands in the court position B6, where we will explain the notation below.

   In certain situations, such as volley, the tennis ball is intercepted by the opponent before it lands. In this case, instead of recording the position where the ball lands in, we record the position where the player intercepts the ball.

   Shot Type Encoding. The shot type reflects the kind of stroke used by the player. The encoding is shown as follows:

   - `s`: serve
   - `f`: forehand ground stroke
   - `b`: backhand ground stroke
   - `v`: volley
   - `c`: slice
   - `o`: overhead (smash)
   - `l`: lob
   - `d`: drop shot

   Extra Notations. In the case where the ball is intercepted by the volleyer, the intended direction of the ball may also be encoded as the character(s) after the shot type encoding, before the ball landing position, e.g., `WbxB6`. This is particularly useful for analysis in the scenario where the ball is intercepted by the opponent. This field is optional as it can be inferred from the landing position of the ball. The encoding is shown as follows, which is optionally charted after the shot type encoding:

   - `i`: down the line
   - `x`: cross court
   - `m`: down the middle

   Note at our main focus of this paper is the spatial distribution of tennis shots and related player movements. Other properties of the game, such as *ball speed*, can be added as part of extra notations in the future, e.g., observational instruments highlighted in Torres-Luque et al. [23].

4. **Player Positions.** The position of the players is an important part of the game, especially in tennis doubles, where the placement of each player is crucial to the strategic game play. And before every shot, the position of the player may change, hence, the player position is recorded along with each shot. To capture this, we use a tuple of the form (X,Y) for singles games and (X, Y, Z, W) for doubles games, where each element represents the court section occupied by the corresponding player right before the time of the shot.
5. **Point Ending.**
   The ending of a point is encoded in the format of `[Winning Player/Team] + [Outcome]`. **Player/Team:** X (Player X in singles games); Y (Player Y in singles games); X (Team X/Y in doubles games); Z (Team Z/W in doubles games).
   Outcome: V (Winner) E (Forced error of opponent); U (Unforced error of opponent).
   For example, `ZV` means Team Z/W wins with a winner point. In the case of forced or unforced error, the type of the error may also be encoded as follows: n (net); w (wide); h (long).

Summary of the overall language

Every point comprises of several shots and players' positions may be changing throughout the game point. For every point, it is encoded as `[Shot Encoding]` followed by optional `[Player Position Encoding]` in chronological order, where the player position refers to the players' positions immediately before the shot is played. At the end of each sequence, the result of the point is noted.

The following is an example charting of a doubles point where player X begins the with a serve, and the team Z/W ultimately wins the point:

```
XsC6, (B0, D4, B6, E9)
WfB6, (A1, C4, B6, D9)
XbD6, (A1, D4, B6, E8)
WfE1, (A0, C4, B6, E9)
ZV
```

4 Strategy Analysis

In this section, we discuss a few analyses enabled by our description language.

4.1 Winning Shot Sequence Identification

Our language effectively describes a tennis game as a sequence of moves in string format. Therefore, similar to chess winning move sequences, such as the Scholar's Mate [6], tennis games also have patterns for winning. When the game is described by strings, such move sequences will be reflected as patterns of regular expressions.

In computer science, there are existing approaches to solving the *common substring of multiple strings* problem. One such algorithm [11] identifies the longest substring that appears in at least k strings. Building on this idea, we adopted the algorithm to identify the substring that is common to the most number of sequences as a winning pattern. We will report a few cases we found in the evaluation section.

Algorithm 1: Grey relational analysis for winning strategy identification

Input: X: Encoded and normalized array of shot sequences where $X[i] = (x_i(1), \ldots, x_i(c))$
n: Total number of shot sequences
c: Number of shots in each sequence
Output: Ranked shot sequences based on average grey relational degree
for $i \leftarrow 1$ **to** n **do**
  $\xi \leftarrow 0.5$;
  let $ref \leftarrow X[i]$;
  // Set reference sequence as $X[i]$
  for $j \leftarrow 1$ **to** n **do**
    for $k \leftarrow 1$ **to** c **do**
      $\Delta_j(k) \leftarrow |ref(k) - x_j(k)|$;
    end
  end
  $M \leftarrow \max_j \max_k \Delta_j(k)$;
  $m \leftarrow \min_j \min_k \Delta_j(k)$;
  for $j \leftarrow 1$ **to** n **do**
    for $k \leftarrow 1$ **to** c **do**
      $\gamma_{0j}(k) \leftarrow \frac{m+\xi \cdot M}{\Delta_j(k)+\xi \cdot M}$;
    end
    $\gamma_{0j} \leftarrow \frac{1}{c}\sum_{k=1}^{c} \gamma_{0j}(k)$;
    // Compute grey relational coefficient between reference sequence and sequence j
  end
  $\gamma_i \leftarrow \frac{1}{n-1}\sum_{\substack{j=1 \\ j \neq i}}^{n} \gamma_{0j}$; // Compute average grey relational degree for reference sequence i excluding self-comparison
end
Sort $(\gamma_i, X[i])$ pairs in descending order based on γ_i;
return *The sorted shot sequences based on* γ_i;

4.2 Grey Relational Analysis

Beyond simple pattern matching, the notation language also enables a deeper analysis of the winning strategy. As our notation language captures matches at a finer granularity and encodes more detailed information, we are able to extract meaningful winning patterns from a much smaller set of data points.

As common winning strategies will have relatively high occurrences in games, the corresponding shot sequence described in TANS will also have a high correlation among the points charted in a game. We employed Grey Relational Analysis (GRA), which measures the similarity between data sequences by comparing the geometric shapes of their curves, allowing us to assess the closeness of their relationships. It is a method well-suited for small sample sizes and capable of revealing significant patterns with a small amount of data.

Data preparation and normalization:

As grey relational analysis mainly deals with numerical data, we first encode each shot in sequences into numerical data using a shot encoding function. This function parses the shot into three components: the shot type, the area of the court, and a numeric spot identifier. These components are mapped to integers using predefined dictionaries. The encoded value of each shot is then calculated using a weighted formula that ensures uniqueness across combinations. Let $X_i = (x_i(1), x_i(2), \ldots, x_i(n))$ be our encoded data sequence, then each data point is mapped to its interval image(D_3 operator) using the following formula [13]:

$$x_i(k)d_3 = \frac{x_i(k) - \min\limits_k x_i(k)}{\max\limits_k x_i(k) - \min\limits_k x_i(k)}; \quad k = 1, 2, \ldots, n \tag{1}$$

Grey relational degree calculation:

Given a sequence $X_0 = (x_0(1), \ldots, x_0(n))$ as the reference sequence and sequences $X_i = (x_i(1), x_i(2), \ldots, x_i(n)), \quad i = 1, 2, \ldots, m$, their grey correlation degree is calculated through the following formula[13]:

$$\gamma(x_0(k), x_i(k)) = \frac{\min\limits_i \min\limits_k |x_0(k) - x_i(k)| + \xi \max\limits_i \max\limits_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \xi \max\limits_i \max\limits_k |x_0(k) - x_i(k)|} \tag{2}$$

$$\gamma(X_0, X_i) = \frac{1}{n} \sum_{k=1}^{n} \gamma(x_0(k), x_i(k)) \tag{3}$$

where $\gamma(\mathbf{X}_0, \mathbf{X}_i)$ is known as Deng's grey relational degree between X_0 and X_i, where ξ is the distinguishing coefficient [10].

Each winning sequence is used as a reference sequence to compute the grey correlation with all the other winning sequences. Then the average of the correlation with all the other winning sequences is taken as a standard to identify the best winning pattern. The summarized algorithm is given in Algorithm 1.

5 Evaluation

Utilizing the language we developed, we charted the full 2022 Australian Open Men's Doubles Final match [2], 2024 Australian Open Men's Doubles Final match [3] and also US Open Women's Doubles Final set [1], excluding points lost due to unforced errors. This resulted in charted data of 111 points. More information about the data and availability is discussed in Appendix A.

With this data, we performed different types of evaluations to show the capability of TANS. First, we show that TANS can be used to carry out traditional tennis analytics; Second, we identify common winning patterns in tennis games; Finally, we apply grey-relational analysis to identify the best winning strategy for a specific match.

5.1 Match Statistics

Traditional match statistics analysis is carried out from manual notations. As TANS does not have data loss, it supports tennis match statistics. For example, when we examined the impact of serving in doubles, we found that 68% of winning points were generated by the serving team in the 2024 Australian Open's match. This indicates the importance of controlling the serving game well. And among these, 37% points started with `*sC6`, which refers to serving down the "T". In addition, we analyzed the opening positions of the serving team that led to winning the point. We identified that 74% of these winning points began with the server positioned at `C4` and the volleyer at `C4`, which actually corresponds to the well-known "I-formation" in tennis doubles strategy.

Moreover, we can use a program to automatically translate TANS into other description languages, such as the notation language used by The Match Charting Project [7]. For example, our charted data of `1. XsD6, (C0, D9) 2. YbcB2, (C4, D9), YF` can be translated into `5b#` in the language defined by TennisAbstract while also providing additional information. It can also be translated to English description: 1st serve down the "T"; backhand return cross court; forced error of opponent.

Shot Sequency	Frequency
vC6	7.4468%
vD5	7.4468%
vB5	6.3830%
oB5	6.3830%
oD5	4.2553%

Table 1: Top 5 winning shots.

5.2 Common Winning Pattern

We focused on ball landing or striking positions during the rallies of each point and extracted them from the charted data. We first analyzed the most common winner shot of both matches. From Table 1, we can see that 7% of the winners are generated by volleying to position C6. While C6 seems to be easy to cover, it often becomes an exposed area in doubles, particularly for teams lacking coordination and on-court synergy. This highlights the tactical advantage of targeting weak spots that emerge from insufficient teamwork. Moreover, around 24% of the winner shots are generated by overhead smash or volley to position B5 or position D5, and this again shows the importance of strong net control in doubles.

We further checked for the most common pattern in the last two or three shots of the winner points charted. From Table 2, the most common two-shot patterns are `C4, B5` and `C4, D5`, each with percentage of around 5%. By symmetry of `B5` and `D5`, we could treat them as variations of one shot where the volleyer volleys cross court. Then this pattern appears in over 10% of the winning patterns. This again highlights the effectiveness of having the net player control the center and execute high-quality

Shot Sequence	Frequency
C4, D5	5.3191%
C4, B5	5.3191%
C4, C6	4.2553%
C4, D6	4.2553%
B3, A8	3.1915%

Table 2: Top 5 Final two-shot placement patterns.

Shot Sequence	Frequency
E6, C4, C6	3.1915%
D6, D4, D5	2.1277%
C6, C4, A6	2.1277%
B6, B4, E5	2.1277%
B6, C4, D5	2.1277%

Table 3: Top 5 Final three-shot placement patterns.

cross-court volleys. Furthermore, from Table 3, the most frequent three-shot winning sequence is E6, C4, C6, which reflects a strategic play: hitting a wide cross-court shot, forcing the opponent to return to the middle, and then the volleyer finished the point with a short, well-placed volley, targeting the same player.

5.3 Winning Strategy from GRA

Sequence	Grey correlation degree
⋆sB6, ⋆bB4, ⋆vE5	0.7626
⋆sB6, ⋆bB4, ⋆vD6	0.7626
⋆sC6, ⋆fE1, ⋆fE6	0.7621
⋆sD6, ⋆fE3, ⋆fD7	0.7603
⋆sB6, ⋆fB4, ⋆vD6	0.7598
⋆sC6, ⋆fC4, ⋆vE5	0.7584
⋆sB6, ⋆fB4, ⋆vD5	0.7584
⋆sD6, ⋆bD4, ⋆oD5	0.7522
⋆sD6, ⋆fD4, ⋆vE5	0.7504
⋆sC6, ⋆lB2, ⋆oD6	0.7429

Table 4: Top 10 winning sequences.

Using the GRA method in Section 4.2, we aim to identify the frequently used strategy for the serving side to win a point. In our evaluation, we calculated the correlation among the winning sequences in the charted men's doubles match. Most of the winner points are short, requiring only three shots, in the case where the point is longer, we take the first three shots for analysis. We also masked the player with ⋆, so that the algorithm can focus on the serving team, regardless of the specific players involved.

(a) sB6: Serve to B6

(b) bB4: Backhand to B4

(c) vE5: Volley to E5

Fig. 3: Screenshot of the winning pattern sB6, bB4, vE5.

The sequences with the top 10 correlation scores are listed in Table 4. Among these, the sequence ⋆sB6, ⋆bB4, ⋆vE5 has the highest average correlation, 0.7626.

This refers to the following strategy: The serving player serving down the "T" towards opponent's backhand. Then returning player returns the ball towards the middle. Positioned near the middle of the net, the volleyer on the serving team intercepts the ball and strikes cross court, resulting in a winning point. The detailed point is shown in Figure 3, using screenshot from the match [3].

Similar strategies were observed in women's doubles tennis. In the final set of the 2019 US Open, the patterns ⋆sC6, ⋆bmC4, ⋆oB6 and ⋆sE6, ⋆bmC4, ⋆oE5 exhibited high correlation scores of 0.8361 and 0.8290, respectively. Both of these patterns involve a powerful serve to the returner's backhand, prompting a return to the volleyer, who then hits it back to the returner.

6 Conclusion

In this paper, we present the Tennis Algebraic Notation System (TANS), a novel tennis notation system inspired by the algebraic notation system of chess. TANS is developed as a formal description to lay the foundation for comprehensive analysis of game strategies. In our evaluation, TANS has been demonstrated to be effective for supporting match statistics analysis, winning pattern discovery and winning strategy identification. As the future work, we aim to leverage the advancement in tennis video analysis, chess analytics, and AI solutions for chess to enable deeper analysis of tennis strategies.

Acknowledgments. We thank the anonymous reviewers for their valuable comments to improve this paper. We dedicate this paper to the Tennis Team of the Peking University Alumni Association (Singapore) for the collaborative spirit and unwavering motivation. This paper is in part supported by the Grant A-8002737-00-00 from the School of Computing, National University of Singapore.

References

1. US open 2019 women's doubles final (2019), https://www.youtube.com/watch?v=xbxPVXOpoA8
2. Australian open 2022 men's doubles final (2022), https://www.youtube.com/watch?v=iZYCl7xq4TI

3. Australian open 2024 men's doubles final (2024), https://www.youtube.com/watch?v=mC_pSNURbzU&t=3933s
4. Mutua madrid open 2025 women's singles final (2025), https://www.youtube.com/watch?v=_O-4_IJFIjQ
5. Chess algebraic notation (Accessed in April 2025), https://www.chess.com/article/view/chess-notation
6. Scholar's mate (Accessed in April 2025), https://www.chess.com/terms/scholars-mate-chess
7. Tennisabstract.com (Accessed in April 2025), https://www.tennisabstract.com/
8. Chiu, C.H., Tsao, S.Y.: Mathematical model for the optimal tennis placement and defense space. International Journal of Sport and Exercise Science **4**(2), 25–36 (2012)
9. (Coordinator), S.J.E.: Portable game notation specification and implementation guide (1994), https://www.saremba.de/chessgml/standards/pgn/pgn-complete.htm
10. Deng, J.: Grey Control Systems. Press of Huazhong University of Science and Technology (1985)
11. Hui, L.C.K.: Color set size problem with applications to string matching. In: Combinatorial Pattern Matching (CPM). Lecture Notes in Computer Science, vol. 644. Springer (1992)
12. Kocib, T., Carboch, J., Cabela, M., Kresta, J.: Tactics in tennis doubles: analysis of the formations used by the serving and receiving teams. International Journal of Physical Education, Fitness and Sports **9**(2), 45–50 (2020)
13. Liu, S.: Grey Systems: Analysis, Methods, Models and Applications. Series on Grey System, Springer, London, second edn. (2016)
14. Liu, Z., Jiang, K., Hou, Z., Lin, Y., Dong, J.S.: Insight analysis for tennis strategy and tactics. In: Proceedings of the 2023 IEEE International Conference on Data Mining (ICDM). pp. 1175–1180. IEEE (2023)
15. Liu, Z., Dong, C., Wang, C., Dong, T.Y., Jiang, K.: Exploring team strategy dynamics in tennis doubles matches. In: Proceedings of the International Symposium on Advances in Computing and Engineering (ISACE) (2024)
16. Martínez-Gallego, R., Crespo, M., Ramón Llin, J., Micó, S., Guzmán Luján, J.F.: Men's doubles professional tennis on hard courts: game structure and point ending characteristics (2020)
17. Martínez-Gallego, R., Ramón-Llin, J., Crespo, M.: A cluster analysis approach to profile men and women's volley positions in professional tennis matches (doubles). Sustainability **13**, 6370 (2021)
18. McIlroy-Young, R., Wang, R., Sen, S., Kleinberg, J., Anderson, A.: Learning models of individual behavior in chess. In: Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. ACM (2022)
19. Nübel, C., Dockhorn, A., Mostaghim, S.: Match point ai: A novel ai framework for evaluating data-driven tennis strategies (2024)
20. Ruoss, A., Delétang, G., Medapati, S., Grau-Moya, J., Wenliang, L.K., Catt, E.: Grandmaster-level chess without search. arXiv preprint arXiv:2304.11415 (2023)
21. Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., Hassabis, D.: Mastering chess and shogi by self-play with a general reinforcement learning algorithm. Nature **550**(7676), 354–359 (2017)
22. Terroba, A., Kosters, W.A., Vis, J.K.: Tactical analysis modeling through data mining: Pattern discovery in racket sports. In: Proceedings of the International Conference on Data Mining (ICDM) (2008)

23. Torres-Luque, G., Fernández-García, A.I., Cabello-Manrique, D., Giménez-Egido, J.M., Ortega-Toro, E.: Design and validation of an observational instrument for the technical-tactical actions in singles tennis. Frontiers in Psychology **9**, 2418 (2018)

A Tennis Game Data Used in Evaluation in TANS

We used TANS format to chart several tennis games for our evaluation, including the game between Bolelli/Vavassori and Bopanna/Ebden in Australian Open 2024 Men's Doubles Final [3], the game between Azarenka/Barty and Mertens/Sabalenka in US Open 2019 Women's Doubles Final [1] and the game between Sabalenka and Gauff in Mutua Madrid Open 2025 Women's Singles Final [4].

The charted data for our evaluation is available at our project's supporting website: `https://tennis-ans.github.io/`. We will use the space to host datasets and other related documents and tools for this project.

B Extension of Tennis Court Partition

An example of fine-grained partition is shown in Figure 4. Column A is subdivided into Column l and Column a; Column C is subdivided into Column p and Column q; c is be used for marking the position along the center line; Column E is subdivided into Column e and Column r; Column B and Column D remain to be Column b and Column d. So the fine-grained columns labels are: l, a, b, p, c, q, d, e, r. This notation is reserved for future computerized processing, which is not used in this paper.

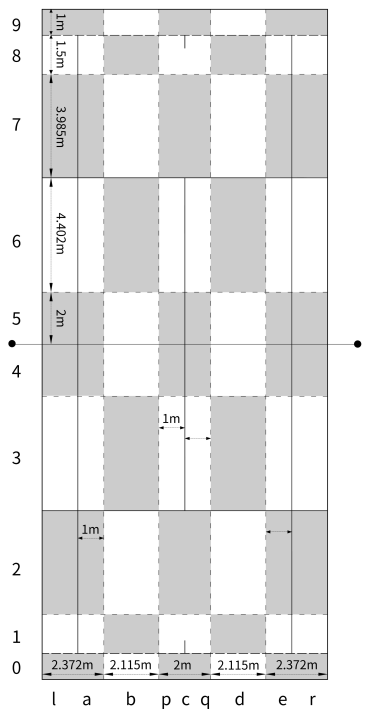

Fig. 4: Extension of Tennis court coordinate in TANS.