

Le système de notation Elo

Quand des joueurs ou des équipes s'affrontent, leurs chances de gagner dépendent de leur force respective. Le système de notation Elo est une méthode de calcul du niveau de compétence relatif des participants à un jeu à somme nulle qui permet d'évaluer l'issue probable d'un match.

Christian Genest
Université McGill

Julien Fageot
École polytechnique
fédérale de Lausanne

Du 24 novembre au 16 décembre 2021 s'est déroulé à Dubaï le match de championnat du monde d'échecs. Il opposait le Norvégien Magnus Carlsen, détenteur du titre depuis 2013, au Russe Ian Nepomniachtchi, vainqueur du tournoi des candidats. Une bourse de 1,2 millions d'euros (soit environ 1,75 millions de dollars canadiens) était destinée au vainqueur.

Le championnat était disputé au meilleur de 14 parties. Une victoire valait 1 point, une défaite 0 point et une partie nulle 1/2 point à chacun des deux concurrents. Au final, Carlsen l'a emporté 7 1/2 à 3 1/2, avec 4 gains, aucun revers et 7 nulles en 11 parties, les trois dernières n'étant pas nécessaires pour couronner le champion. Carlsen a ainsi obtenu son 5^e titre mondial consécutif.

Ce résultat est conforme au système de notation Elo que la Fédération internationale des échecs (FIDE) utilise depuis 1970 pour coter les joueurs en fonction de leur niveau de compétence relatif. En effet, Carlsen était classé au départ premier joueur mondial avec une cote Elo de 2856, alors que son adversaire, meilleur joueur russe, était classé 5^e avec une cote de 2782.

Lors d'une partie entre des joueurs A et B dont les cotes sont θ_A et θ_B , le score moyen de A prédit par la méthode Elo est

$$s_{AB} = \frac{1}{1 + 10^{-(\theta_A - \theta_B)/400}}, \quad (1)$$

sachant que A marquera un point s'il gagne, aucun s'il perd, et 1/2 point en cas de partie nulle. Par voie de conséquence, le score moyen de B est simplement $s_{BA} = 1 - s_{AB}$.

Au début du championnat, on avait $\theta_A = 2856$ pour Carlsen et $\theta_B = 2782$ pour Nepomniachtchi, d'où $s_{AB} \approx 0,60$, laissant présager

que le Norvégien prendrait éventuellement l'ascendant sur son adversaire puisque a priori, il marquerait en moyenne 8 1/2 points en 14 parties et que le Russe n'en ferait que 5 1/2.

Fondement mathématique de la méthode

Pour déterminer le score moyen s_{AB} d'un joueur A contre un joueur B , une méthode naïve consisterait à prendre la moyenne des points marqués par le premier dans les parties l'opposant au second. Or, dans un jeu comme les échecs, la vaste majorité des centaines de milliers de joueurs ne se sont jamais rencontrés et même lorsqu'ils l'ont fait, le nombre de parties qu'ils ont disputées est presque toujours trop petit pour une estimation de qualité. Le système Elo pallie cette difficulté en s'appuyant sur un modèle mathématique permettant de classer les joueurs et, par là même, d'estimer le résultat escompté lors de leur affrontement.

L'expression (1) est inspirée d'un modèle conçu dans les années 1920 par le mathématicien allemand Ernst Zermelo et raffiné dans les années 1950 par le statisticien canadien Ralph Bradley et son collègue Milton Terry. Dans ces travaux et dans la suite de cet article, on suppose que toute partie fait un gagnant et un perdant. Il s'agit donc d'un modèle simplifié, en particulier aux échecs où la nulle est le résultat le plus souvent observé entre joueurs de l'élite mondiale.

Lorsque toute partie se solde par une victoire (1) ou une défaite (0), le score moyen s'interprète alors comme la probabilité que le joueur A l'emporte sur le joueur B , que l'on note p_{AB} . Le modèle de Bradley-Terry-Elo suppose que cette probabilité peut s'exprimer sous la forme

$$p_{AB} = G(\theta_A - \theta_B), \quad (2)$$

où $G: \mathbb{R} \rightarrow (0, 1)$ est une fonction continue et strictement croissante telle que $G(-x) + G(x) = 1$ pour tout réel $x \in \mathbb{R}$ et $G(x) \rightarrow 1$ quand $x \rightarrow \infty$. Un cas particulier d'une telle fonction est illustré à la figure 1. On retrouve ainsi que

$$p_{AB} + p_{BA} = G(\theta_A - \theta_B) + G(\theta_B - \theta_A) = 1,$$

les nulles étant exclues. Puisque $G(0) = 1/2$, deux adversaires de même calibre ont ainsi la même probabilité a priori de l'emporter ; plus $\theta_A - \theta_B$ augmente, plus une victoire de A sur B a des chances de se produire, jusqu'à devenir certaine à la limite.

Le choix le plus commun de G est la loi logistique de paramètre $\beta > 0$, définie en tout réel $x \in \mathbb{R}$ par l'expression

$$L(x) = 1/(1 + e^{-\beta x}). \quad (3)$$

En prenant $G = L$ et $\beta = \ln(10)/400$, on tombe sur la formule (1) employée par la FIDE. Dans ses travaux initiaux, le physicien hongro-américain Arpad Elo, qui a promu cette modélisation dans le monde des échecs, avait plutôt privilégié l'emploi d'une fonction de répartition normale centrée mais non réduite pour G , ce qui mène toutefois à une expression moins explicite pour s_{AB} .

Propriétés du modèle probabiliste

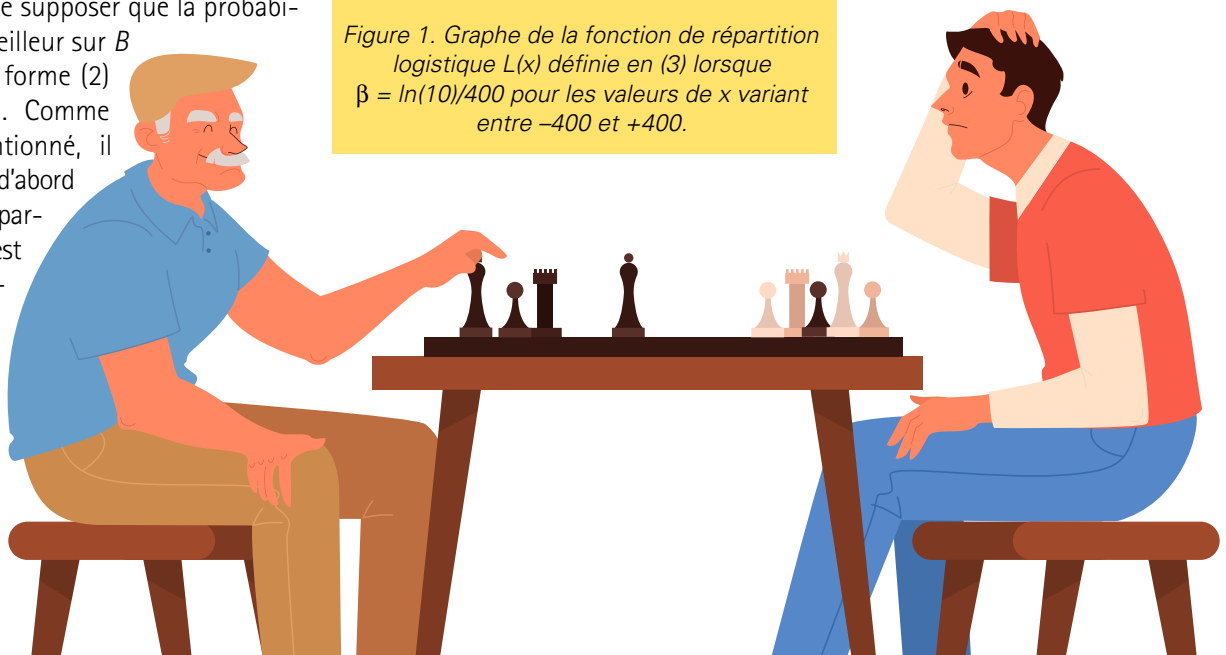
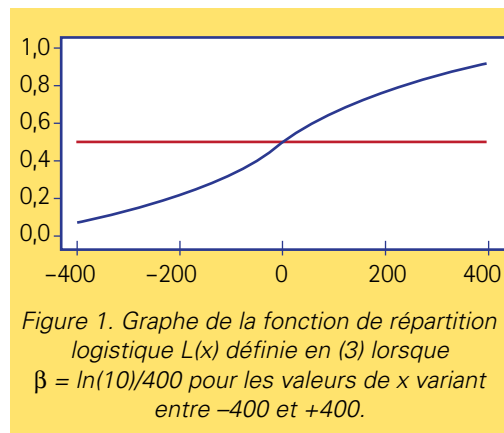
Le modèle de Bradley-Terry-Elo est avantageux en ce qu'il permet de classer tous les concurrents à un jeu en fonction de leur force relative, même s'ils ne se sont jamais affrontés directement. On peut aussi s'en servir pour faire de la prévision, notamment au moyen de simulations.

Toutefois, le fait de supposer que la probabilité que A ait le meilleur sur B s'exprime sous la forme (2) n'est pas anodin. Comme on l'a déjà mentionné, il sous-entend tout d'abord qu'il n'y a pas de parties nulles. Si tel n'est pas le cas, des ajustements doivent être apportés au modèle.

L'expression (2) présuppose aussi que l'issue d'un affrontement entre A et B ne dépend que de leur force relative, $\theta_A - \theta_B$, mesurée sur une échelle linéaire. En particulier, le fait de jouer en premier (comme les blancs aux échecs), d'être plus reposé que son adversaire ou d'évoluer devant ses partisans n'influe pas sur le résultat (ce qui est contestable dans le sport professionnel).

De plus, il est important de comprendre que la cote Elo d'un joueur n'a aucune valeur intrinsèque : elle ne s'interprète qu'en fonction de celle de ses adversaires puisque l'on peut à loisir translater toutes les cotes par la même constante sans affecter les différences. De même, toutes les cotes peuvent être multipliées par une constante $c > 0$ au choix sans changer les probabilités de gain, à condition de remplacer $G(x)$ par $G(x/c)$.

Dans le cas de la formule (1), tracée à la figure 1 pour des valeurs de $\theta_A - \theta_B$ entre -400 et +400, le choix du paramètre β a été conçu pour qu'un écart de 100 points entre deux joueurs A et B donne au meilleur des deux environ 64% des chances de gain. Si l'écart est de 200 points, la probabilité passe à 75% et elle est d'environ 91% quand l'écart est de 400 points.



Aux échecs, la cote de la grande majorité des joueurs classés par la FIDE varie entre 1000 et 3000. La distribution des cotes est généralement asymétrique et varie aussi beaucoup de pays en pays. Par ailleurs, il existe plusieurs variantes du système et celui-ci a en outre été adapté à d'autres jeux et sports, notamment le go et le scrabble, divers jeux en ligne, mais aussi le football, le tennis et – on le verra plus loin – le hockey sur glace.

Actualisation des cotes

La force relative des concurrents à un jeu fluctue bien sûr au fil du temps. Certains d'entre eux s'améliorent à mesure qu'ils prennent de l'expérience ; d'autres faiblissent avec l'âge ou, dans le cas des équipes sportives, en raison de blessures, de retraites ou d'échanges défavorables.

Pour tenir compte de cette dynamique, la méthode Elo propose une règle de mise à jour des cotes après chaque partie. Appelons $\theta_{A,n}$ et $\theta_{B,n}$ les cotes des joueurs A et B au temps n où ils s'apprentent à s'affronter. Elo a proposé qu'à l'issue de la rencontre, la différence $\theta_{A,n} - \theta_{B,n}$ de leurs forces relatives leur soit ré-attribuée proportionnellement en fonction du « degré de surprise » du résultat et d'un nombre réel positif k appelé coefficient de développement.

Ainsi, appelons s le score de la partie opposant A à B avec $s = 1$ en cas de victoire de A , $s = 0$ en cas de défaite aux mains de B et $s = 1/2$ pour une nulle. La nouvelle cote Elo de A sera alors

$$\theta_{A,n+1} = \theta_{A,n} + k(s - s_{AB,n}),$$

où $s_{AB,n}$ est le score attendu calculé selon la formule (1) à partir des cotes $\theta_{A,n}$ et $\theta_{B,n}$. On vérifie aisément que cette cote augmente lorsque $s = 1$, et ce d'autant plus que $s_{AB,n}$ est petit. Ainsi, une victoire de A contre un joueur de grande force est mieux récompensée. De même, la cote de A diminue en cas de défaite, soit quand $s = 0$. Enfin, la cote de A est rehaussée en cas de partie nulle à la condition que le joueur B ait une meilleure cote initiale, soit lorsque $s_{AB,n} < 1/2$.

Pour les joueurs professionnels, la FIDE a fixé $k = 10$, de sorte que durant le championnat du monde, chaque victoire de Carlsen lui a rapporté $10 \times (1 - 0,60) \approx 4$ points sur l'échelle Elo, tandis que chaque partie nulle lui faisait perdre $10 \times (0,5 - 0,60) \approx 1$ point. Après le championnat, sa cote est donc passée à

$$2856 + 4 \times 4,0 - 7 \times 1,0 = 2865,$$

tandis que celle de Nepomniachtchi a chuté de 2782 à 2773.

D'autres valeurs de k sont utilisées en fonction du niveau des joueurs. Par exemple, on prend $k = 40$ pour les nouveaux joueurs jusqu'à leur trentième partie ou pour les jeunes joueurs avant leurs 18 ans, ce qui leur permet de progresser rapidement jusqu'à leur niveau réel. Puis k est fixé à 20 tant que leur classement reste inférieur à 2400 points.

Propriétés de la méthode d'actualisation

Dans le cas particulier où toutes les parties font un vainqueur, on voit que de façon générale, la méthode de mise à jour proposée par Elo fait intervenir une fonction $M: \mathbb{R} \rightarrow (0, \infty)$ continue et strictement décroissante telle que $M(x) \rightarrow 0$ quand $x \rightarrow \infty$. Cette fonction permet d'actualiser la cote des joueurs A et B comme suit :

Si A bat B , alors

$$\begin{aligned} \theta_{A,n+1} &= \theta_{A,n} + M(\theta_{A,n} - \theta_{B,n}), \\ \theta_{B,n+1} &= \theta_{B,n} - M(\theta_{A,n} - \theta_{B,n}). \end{aligned}$$

Si B bat A , alors

$$\begin{aligned} \theta_{A,n+1} &= \theta_{A,n} - M(\theta_{B,n} - \theta_{A,n}), \\ \theta_{B,n+1} &= \theta_{B,n} + M(\theta_{B,n} - \theta_{A,n}). \end{aligned}$$

Pour la méthode employée par la FIDE, on a $M(x) = kL(-x)$ pour tout réel $x \in \mathbb{R}$.

Cette formule de mise à jour a une interprétation probabiliste heuristique. En effet, soient θ_A et θ_B les forces relatives réelles mais inconnues des joueurs A et B . Supposons que $G(\theta_A - \theta_B)$ soit la véritable probabilité que A l'emporte sur B . La valeur espérée du changement de cote du joueur A au temps n est alors donnée par l'expression suivante :

$$\begin{aligned} &M(\theta_{A,n} - \theta_{B,n})G(\theta_A - \theta_B) \\ &- M(\theta_{B,n} - \theta_{A,n})G(\theta_B - \theta_A). \end{aligned}$$

Si la méthode d'actualisation est bonne, on s'attendrait à ce que cette espérance soit nulle si $\theta_{A,n} - \theta_{B,n} = \theta_B - \theta_A$. Sachant que la fonction M est décroissante, on souhaiterait aussi qu'à l'issue de la partie opposant A à B , la différence $(\theta_{A,n+1} - \theta_{B,n+1}) - (\theta_A - \theta_B)$ soit plus près de zéro que $(\theta_{A,n} - \theta_{B,n}) - (\theta_A - \theta_B)$, qui était sa valeur avant leur affrontement.

Le même raisonnement vaut pour le joueur B et on peut vérifier que ces attentes seront comblées si les fonctions G et M satisfont, pour tout réel $x \in \mathbb{R}$,

$$M(x)/M(-x) = G(-x)/G(x). \quad (4)$$

Partant de cette équation d'équilibre, on peut aussi montrer, au prix de certains efforts, que pour une fonction M préétablie, la seule solution G possible est donnée, en tout $x \in \mathbb{R}$, par

$$G(x) = \frac{M(-x)}{M(x) + M(-x)}.$$

La règle d'actualisation spécifiquement proposée par Elo sous-tend donc au départ une probabilité de gain logistique, c'est-à-dire le fait de prendre $G = L$.

Notons ici une subtilité : si on commence par fixer la fonction G , il existe une infinité de fonctions M obéissant à la relation

(4). En effet, quelle que soit la constante $k > 0$, on obtient déjà une solution en posant $M(x) = kG(-x)$ pour tout réel $x \in \mathbb{R}$. Toutefois, on peut aussi prendre $M(x) = kG(-x)S(|x|)$, où S est une fonction telle que $S(0) = 1$ et $S(x) = S(-x)$ pour tout réel $x \in \mathbb{R}$. Il faut cependant s'assurer que le choix de S garantit la décroissance de la fonction M .

Un résultat de convergence

Il est possible de faire un rapprochement entre la méthode d'actualisation de Elo et la méthode itérative du gradient stochastique employée en apprentissage automatique dans le cadre d'un modèle de régression logistique.

Imaginons qu'il y ait au total N joueurs et que le vecteur $(\theta_1, \dots, \theta_N)$ de leurs forces respectives soit inconnu mais fixe. Imaginons aussi qu'au départ, on dispose de cotes approximatives pour ces concurrents, disons $(\theta_{1,0}, \dots, \theta_{N,0})$, et qu'une suite infinie de parties soient disputées, chacun d'elles opposant deux individus choisis uniformément au hasard. Enfin, imaginons qu'après chaque rencontre, le vecteur des forces respectives soit actualisé par la méthode Elo en modifiant ses composantes i et j correspondant aux deux joueurs qui se sont affrontés.



Arpad Elo

Né le 25 août 1903 à Egyházaskesző, en Autriche-Hongrie, Arpad Elo (Élő Árpád en hongrois) a émigré aux États-Unis en 1913 avec ses parents

d'origine modeste. Après avoir étudié la physique à l'Université de Chicago, il a enseigné à l'Université Marquette de Milwaukee, au Wisconsin.

Joueur d'échecs doué, Elo a remporté huit fois le championnat d'échecs de son état et a présidé la Fédération américaine des échecs de 1935 à 1937. Il est mondialement connu pour la mise au point du système de notation éponyme, qu'il a conçu dans les années 1950 après que des critiques aient été formulées à l'encontre d'un premier système développé par Kenneth Harkness.

La nouvelle approche a été adoptée par la Fédération internationale des échecs (FIDE) en 1970 et Elo a été fait membre honoraire de cet organisme en 1981.

Une première analyse détaillée du système de notation Elo (parfois écrit à tort en majuscules, ELO, comme s'il s'agissait d'un acronyme) a été proposée par Elo lui-même dans son ouvrage paru en 1978, *The Rating of Chessplayers, Past and Present*. Sa méthode fait aujourd'hui partie intégrante du monde des échecs. Elle sert à déterminer comment s'affrontent les joueurs dans les opens d'échecs du monde entier et chaque joueur suit scrupuleusement son évolution au classement qui mesure ses progrès et le positionne dans la hiérarchie échiquéenne.

Arpad Elo est décédé le 5 novembre 1992 (à 89 ans) à Brookfield (Wisconsin).

En supposant que les fonctions G et M obéissent aux conditions déjà mentionnées, et en supposant de plus que la première dérivée de M est bornée en valeur absolue, le probabiliste britannique David Aldous a démontré que la suite de vecteurs de terme $(\theta_{1,n}, \dots, \theta_{N,n})$ atteindra éventuellement, quand n tendra vers l'infini, une distribution stationnaire. La question de savoir si cette loi stationnaire constitue ou non une bonne approximation de $(\theta_1, \dots, \theta_N)$ reste théoriquement ouverte, bien qu'elle soit confirmée par de nombreuses simulations et les dizaines d'année d'expérience des joueurs d'échecs eux-mêmes dans l'usage du classement Elo.

Cependant, bien qu'un résultat de convergence comme celui d'Aldous puisse inspirer confiance dans la méthode d'Elo, il est peu probable qu'il existe un vecteur $(\theta_1, \dots, \theta_N)$ sous-jacent d'aptitudes réelles des concurrents. Après tout, les joueurs évoluent au fil du temps et les équipes sportives sont en constante transformation. Le vecteur des niveaux réels, s'il existe, est donc dynamique.

Ce qui fait le succès de la méthode d'Elo, c'est son habileté à suivre cette évolution sans pour autant s'appuyer sur un modèle stochastique bien précis. C'est un fait avéré dans le domaine des échecs, mais aussi dans de nombreux autres contextes où, bien sûr, le choix du paramètre β et celui du coefficient de développement k doivent être adaptés.

Le système Elo au hockey

Pour illustrer la méthode Elo et sa capacité prédictive dans un contexte plus familier aux lecteurs d'Accromath que celui des échecs, considérons brièvement son utilisation dans le domaine du hockey sur glace, où diverses adaptations ont été élaborées pour exploiter les résultats de tous les matchs réguliers et éliminatoires disputés dans la Ligue nationale de hockey (LNH) depuis sa création à Montréal en 1917. Ces données sont stockées sur le site hockey-reference.com.

À titre d'exemple, les analystes Ryan Best et Neil Paine du site Web américain *FiveThirtyEight* ont proposé d'attribuer un classement Elo qui tient compte des paramètres suivants :

- a) Chaque équipe se voit attribuer initialement une cote Elo de 1380.
- b) La règle de mise à jour utilise un facteur de développement de $k = 6$.
- c) On tient compte de « l'avantage de la glace » de sorte qu'à cotes égales, le club qui joue à domicile a 57,1 % de l'emporter, plutôt que 50 %.
- d) L'ajustement de cotes est majoré de 25 % lors des matchs éliminatoires.
- e) En début de saison, la cote de chaque équipe est révisée par la formule

$$0,7 \times \text{cote finale de la saison précédente} + 0,3 \times 1505.$$

Ainsi, la cote Elo moyenne de la ligue oscille bon an mal an autour de 1500.

En outre, le modèle de Best et Paine tient compte de la différence de buts de chaque partie et d'une dépendance temporelle (appelée autocorrélation) entre les matchs successifs.

La figure 2 montre l'évolution de la cote Elo des Canadiens de Montréal fondée sur le modèle de Best et Paine. Le pic de la courbe se situe dans la seconde moitié des années 1970, époque pendant laquelle l'équipe a remporté quatre Coupes Stanley d'affilée.

On observe d'autres sommets dans la seconde moitié des années 1940 et au milieu des années 1990, ainsi qu'un long plateau d'excellence à la fin des années 1950, alors que l'équipe a raflé cinq championnats

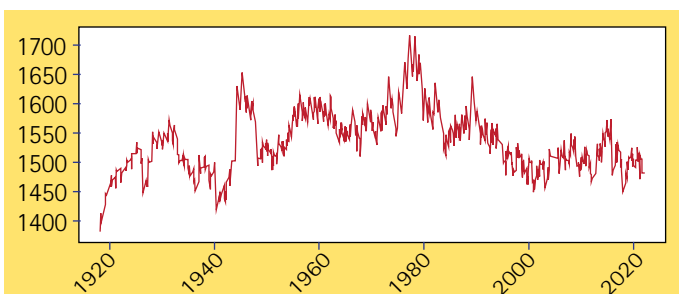


Figure 2. Évolution de la cote Elo des Canadiens de Montréal depuis la saison 1917-18.

consécutifs (1956 à 1960). En revanche, la cote Elo du club oscille autour de la moyenne depuis le début des années 2000. Sa participation à la finale de la Coupe Stanley en 2021 constitue donc un exploit qui a, hélas !, peu de chances de se reproduire dans un proche avenir.

Les cotes Elo étant comparatives, il est plus instructif encore de comparer leur évolution pour différents clubs. C'est ce qui est fait à la figure 3 pour les Canadiens de Montréal (en rouge), les Maple Leafs de Toronto (en bleu) et les cinq équipes qui ont remporté la Coupe Stanley entre 2008 et 2017, soit les Bruins de Boston (orange), les Red Wings de Détroit (vert), les Black Hawks de Chicago (noir), les Kings de Los Angeles (mauve) et les Penguins de Pittsburgh (jaune).

Dans la figure 3, le club champion est indiqué par un cercle. En 2018, la coupe a été remportée par les Capitals de Washington, dont la courbe Elo n'est pas incluse. Comme la figure le montre bien, ce n'est pas toujours la meilleure équipe qui l'emporte en finale¹.

1. Voir l'article intitulé « Que le meilleur gagne ! » dans le numéro d'hiver/printemps 2020 (vol. 15, no 1) d'Accromath.

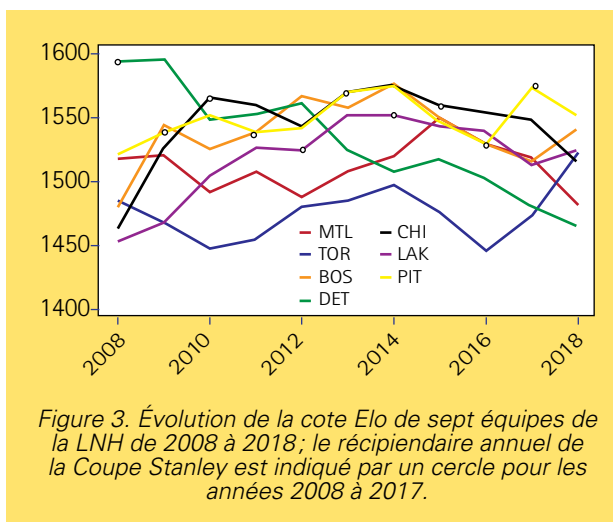


Figure 3. Évolution de la cote Elo de sept équipes de la LNH de 2008 à 2018; le récipiendaire annuel de la Coupe Stanley est indiqué par un cercle pour les années 2008 à 2017.

Conclusion

Élément incontournable du monde des échecs mais s'appliquant au cas bien plus général d'un jeu à somme nulle entre deux joueurs, le système de notation Elo permet d'évaluer la force relative des concurrents même dans le cas où ceux-ci ne se sont jamais affrontés. Ses mises à jour successives au fur et à mesure des parties constituent une façon simple et efficace de capturer la dynamique d'évolution des joueurs ou des équipes, comme on l'a illustré dans le cas de la LNH.

La cote Elo des joueurs d'échecs

Un débutant a aujourd'hui une cote Elo d'environ 1000, là où 2000 constitue le classement d'un bon joueur amateur. Les membres du groupe sélect des grands maîtres internationaux doivent avoir atteint au moins un Elo de 2500. Ce groupe coïncide plus ou moins avec les joueurs professionnels et compte aujourd'hui plus de 1000 joueurs dans le monde. À peine une quinzaine de joueurs sont parvenus jusqu'ici à dépasser la barre symbolique des 2800 Elo dans leur carrière.

Les logiciels d'échecs ont aussi une cote Elo. Le meilleur programme actuel, *Stockfish*, a un classement d'environ 3700, ce qui montre à quel point la machine surpasse l'humain aux échecs : la valeur de p_{AB} entre *Stockfish* (A) et Magnus Carlsen (B) avoisine 0,992... (soit une moyenne de 992 victoires et 8 défaites pour l'ordinateur en 1000 parties, sans match nul). C'est le même écart qu'entre Carlsen et un bon joueur amateur de 2000 Elo, un niveau atteint par 30 000 joueurs.

On peut aussi se demander si le classement Elo peut servir à comparer les joueurs de différentes époques. Par exemple, est-ce qu'un joueur évalué à 2700 Elo dans les années 1980 a un niveau comparable à un joueur de 2700 aujourd'hui ? La question est complexe, notamment car le jeu lui-même a évolué. Par exemple, les joueurs, des professionnels jusqu'aux amateurs les plus motivés, ont intégré les logiciels d'échecs dans leur pratique. Mais des phénomènes d'inflation ont aussi été observés qui sont induits par la façon même dont le calcul du Elo est effectué. Ainsi, au début des années 2000, le classement moyen du Top 50 a augmenté de quelques dizaines de points, ce qui fait qu'un classement de 2700 aurait vraisemblablement moins de valeur en 2010 qu'en l'an 2000.