

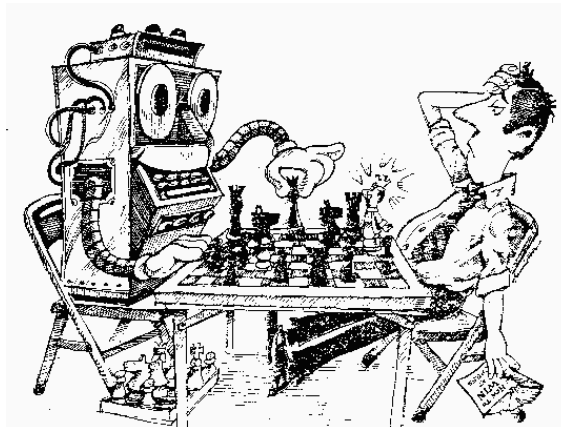
Računalniška primerjava svetovnih prvakov

Matej Guid in Ivan Bratko

Laboratorij za umetno inteligenco, Fakulteta za računalništvo in informatiko, Univerza v Ljubljani

Kdo je bil najboljši šahist vseh časov? To vprašanje že od nekdaj buri duhove svetovne šahovske javnosti, vendar dobro utemeljenega, objektivnega odgovora ni, saj je potrebna primerjava med igralci, ki so živeli v različnih obdobjih in se nikoli niso mogli pomeriti med seboj za šahovnico. Pojavitev vse močnejših šahovskih programov omogoča objektivnejše odgovore na to vprašanje. Vendar pa so kljub temu v dosedanjih tovrstnih raziskavah računalnike v glavnem uporabljali le za statistične obdelave *rezultatov* šahovskih partij. Le-ti pa ne izražajo vedno dejanske šahovske moči igralcev, še posebej ker je kvaliteta igre skozi desetletja na splošno močno napredovala. Jeff Sonas je leta 2005 izvedel obsežno raziskavo, ko je s pomočjo dobro premišljenega ratingovega sistema primerjal rezultatsko uspešnost igralcev vse od leta 1840 naprej, vendar bi zelo težko kdo verjel, da bi prvega uradnega svetovnega prvaka, Wilhelma Steinitza, ki po rezultatih Sonasove raziskave sodi med deset najboljših igralcev vseh časov, dejansko bilo mogoče uvrstiti tako visoko po njegovi šahovski moči, odraženi z njegovimi potezami na šahovnici.

Naš pristop je bil drugačen: zanimala nas je *kvaliteta igre* šahistov, ki smo jo ocenjevali s pomočjo računalniške analize posameznih *potez*. Osnovali smo tudi metodo za določanje kompleksnosti pozicij, da bi pri ocenjevanju upoštevali razlike med različnimi stili igranja. Mirni *pozicijski igralci* so namreč v svojih tipičnih partijah navadno imeli manj priložnosti za grobe taktične napake kot *taktični igralci*, ki so ravno z ustvarjanjem kompleksnejših pozicij postavljali svoje nasprotnike pred nerešljive probleme, pa čeprav tudi sami pri tem niso ostali nezmotljivi. Prav tako podajamo skrbno izbrano metodologijo za uporabo računalniških šahovskih programov za ocenjevanje kvalitete igre šahistov. Gre seveda le za en (vendar zelo pomemben) vidik merjenja šahovske moči, ki ne upošteva psiholoških in ostalih dejavnikov, ki so prav tako prisotni v šahovski igri. Četudi niti računalniki niti ljudje nikoli ne bodo mogli podati dokončnega odgovora na vprašanje, kdo je bil zares najboljši šahist vseh časov (za to enostavno ni zabeleženih dovolj podatkov, na voljo imamo praktično le odigrane poteze in rezultate partij, pa tudi kriteriji za primerjavo bodo stvar diskusije), so nekatere metode vendarle bolj objektivne od drugih in računalniki kljub vsemu lahko podajo marsikatero koristne odgovore, če jih le uporabimo na primeren način.



Metode ocenjevanja

Ocenjevali smo štirinajst svetovnih prvakov, ki so vladali na svetovnem šahovskem prestolu od prvega uradnega dvoboja za svetovnega prvaka leta 1886 pa vse do leta 2006. Obravnavali smo partije njihovih neposrednih dvobojev za naslov šahovskega prvaka, bodisi so bili v njih v vlogi izzivalca bodisi v vlogi branitelja naslova. Osnovali smo več kriterijev za ocenjevanje. Osnovni kriterij je povprečno odstopanje med ocenami potez, ki so jih svetovni prvaki odigrali v obravnavanih partijah, in med potezami, ki jih je v istih pozicijah predlagal računalnik. Prav tako smo izračunali povprečno število grobih napak in performanse igralcev v primeru, da bi vsi imeli opravka z enako kompleksnimi pozicijami. Naše analize med drugim tudi jasno kažejo, da je odstotek najboljših potez odvisen od pozicije same in da je zelo močno povezan z razliko med ocenama najboljših dveh potez (kot jih predlaga računalnik): večja kot je razlika med najboljšima potezama, lažje je najti najboljšo potezo. Z opazovanjem povprečnega seštevka materiala na šahovnici tekom partije pa smo želeli ugotoviti težnje posameznih igralcev po poenostavitvah pozicije.

Ocenjevanje vsake partije se je pričelo pri 12. potezi. Današnji računalniški programi namreč slabo ocenjujejo poteze v začetni fazi igre, hkrati pa bi zgodnejši pričetek ocenjevanja partij bil po vsej verjetnosti v prid svetovnim prvakom mlajših generacij, ki so bili v fazi otvoritve zaradi stalnega napredka teorije otvoritev boljše pripravljene od njihovih predhodnikov. Vendar pa tudi ne gre prezreti dejstva, da so slednji imeli na voljo več časa za razmišljanje. Kasnejši pričetek ocenjevanja bi po drugi strani vodil do izgube koristnih informacij in 12. poteza je bila izbrana kot ustrezen kompromis.

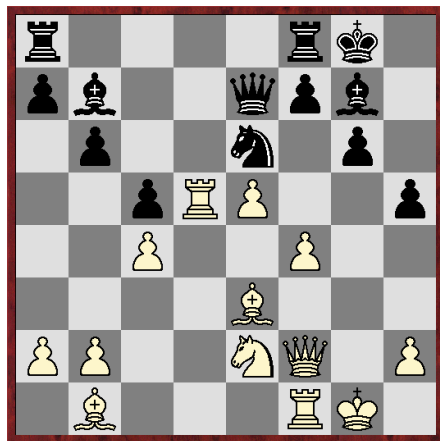
Pri izračunu nismo upoštevali potez, pri katerih imata tako najboljše ocenjena poteza kot odigrana poteza oceno izven območja od -2.00 do +2.00. Namreč ko šahist ocenjuje, da je njegova pozicija *dobljena* oz. da ima prednost, ki zadošča, da partijo lahko odloči v svojo korist tudi ob morebitni najboljši igri nasprotnika, včasih namenoma ne povleče najboljše poteze, ampak se odloči za potezo, ki je »dovolj dobra« in vodi do zmage z manj tveganja. Prav tako šahisti v izgubljenih pozicijah, kjer ocenijo, da se niti z morebitnimi najboljšimi potezami ne morejo rešiti poraza, če nasprotnik ne naredi kakšne večje napake, včasih namenoma povlečejo poteze, za katere sicer vedo, da so objektivno slabše, vendar jim dajejo večje *praktične možnosti*, da se rešijo iz težkega položaja. Neprimerno bi bilo obsojati igralce, ko so na ta povsem legitimen način poskušali priti oz. tudi prišli do zelenega rezultata.

Verodostojnost programa za ocenjevanje

Za ocenjevanje je bil uporabljen šahovski program Crafty. Potrebovali smo program odprte kode, saj so za potrebe analize bile potrebne nekatere manjše spremembe programa. Crafty sicer ima moč šahovskega velemejstra, kljub temu pa se bo prav gotovo marsikdo vprašal, kako je mogoče za ocenjevanje uporabiti program, ki je po vsej verjetnosti slabši vsaj od večine obravnavanih svetovnih prvakov. Vendar, ocenjenih je bilo skupno več kot 37.000 pozicij in čeprav Craftyjeve ocene niso vselej pravilne, morajo za tovrstno analizo biti le dovolj natančne, da se občasne manjše napake v ocenah statistično izničijo. Matematični model in nadaljnje raziskave kažejo, morda nekoliko presenetljivo, da za uspešno rangiranje igralcev po uporabljeni metodologiji ni potrebno, da je izbrani program boljši od ocenjevanih igralcev (čeprav je za čimbolj zanesljive rezultate priporočljivo, da je tem boljši) in da bi uporaba katerega od močnejših šahovskih programov po vsej verjetnosti vodila do podobnih rezultatov.

Številni argumenti govorijo v prid računalniškim programom kot ocenjevalcem. Pri ocenjevanju pozicij navajajo numerične ocene, ki so za naš namen bile neprimerno bolj uporabne kot tipične ocene ljudi šahistov, hkrati pa se pri ocenjevanju ves čas držijo istih pravil in so zato pri ocenah veliko bolj konsistentni od ljudi in tudi niso pod vplivom čustev in drugih, za ocenjevanje motečih dejavnikov, ter so povrh vsega izredno dobri pri detektiranju taktičnih napak. Poleg tega je bil Crafty pri ocenjevanju omejen na fiksno globino preiskovanja (in ne časovno), kar pomeni, da je bolj kompleksnim pozicijam avtomatsko namenjal več časa.

Vsi vemo, da različni šahovski programi podajajo različne ocene za iste ocenjevane pozicije in splošna intuicija nas kaj lahko vodi do zaključka, da bi ocenjevanje z različnimi programi dalo povsem različne rezultate. Pa je res tako? Poglejmo si pozicijo na naslednjem diagramu:



Botvinnik-Talj, 17.partija za naslov svetovnega prvaka, Moskva 1961

V poziciji na diagramu je Talj odigral potezo 23...Sc7! in kasneje, v 83. potezi, tudi zmagal to partijo. Poglejmo še Crafty-jeve ocene te pozicije pri različnih globinah iskanja:

Globina	Najboljša poteza	Ocena
2	Lxd5	-1.46
3	Lxd5	-1.44
4	Lxd5	-0.75
5	Lxd5	-1.00
6	Lxd5	-0.60
7	Lxd5	-0.76
8	Tad8	-0.26
9	Lxd5	-0.48
10	Tfe8	-0.14
11	Lxd5	-0.35
12	Sc7	-0.07

Ocene Crafty-ja za zgornjo pozicijo pri različnih globinah iskanja

Kot zgornja tabela nazorno prikazuje, Crafty pri različnih globinah iskanja pozicijo na diagramu ocenjuje zelo različno: ne samo, da se izrazito menjajo ocene (po njegovih izračunih) najboljše poteze v poziciji, tudi sama najboljša poteza je od globine 7 naprej vselej različna kot pri prejšnji globini. Podobno velja tudi za ostale pozicije: ocene in predlagane najboljše poteze se pogosto spreminjajo z globino iskanja (kako močno je seveda odvisno od ocenjevanih pozicij).

Pojavi se vprašanje: »**Kako bi Crafty razvrstil ocenjevane svetovne prvake, če bi ga omejili na različne globine iskanja?**« Glede na omenjene razlike v ocenah pri različnih globinah, nam intuicija sugerira na videz jasen odgovor: pri vsaki globini bi dobili povsem različne rezultate. Vendar pa so naše nadaljnje raziskave pokazale, da bi Crafty tudi pri preiskovanju do različnih globin ohranil rangiranje igralcev. Prav pri vseh globinah sta Capablanca in Kramnik na prvih dveh mestih, Euwe in Steinitz pa na zadnjih dveh. Pri nekaterih igralcih vmes se vrstni red sicer minimalno spreminja, kar je posledica dejstva, da so njihovi rezultati zelo izenačeni.

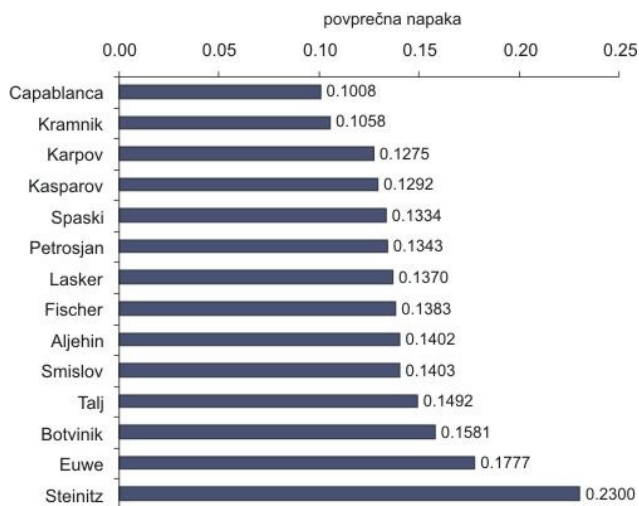
Kaj ta rezultat pomeni? Leta 1982 je Ken Thompson v znanstvenem članku »Computer chess strength« primerjal med seboj programe pri različnih globinah iskanja. Po njegovih rezultatih je program, ki išče samo eno polpotezo globlje, boljši za več kot 200 ratingovskih točk od enakega programa z (za eno polpotezo) nižjo globino iskanja. Naši rezultati ocenjevanja s Crafty'jem do različnih globin torej kažejo na to, da **čepprav bi uporabili različno močne verzije Crafty-ja, bi še vedno dobili praktično enako rangiranje igralcev**. Povsem umestno je tudi sklepanje, da bi Crafty na zelo podoben način ohranil vrstni red igralcev tudi pri globinah iskanja do 20 in še dlje! Crafty pri tako velikih globinah iskanja pa bi se po vsej verjetnosti lahko povsem enakomerno kosal z najboljšimi programi, ki bi preiskovali do nekoliko nižjih globin.

Rezultati primerjave

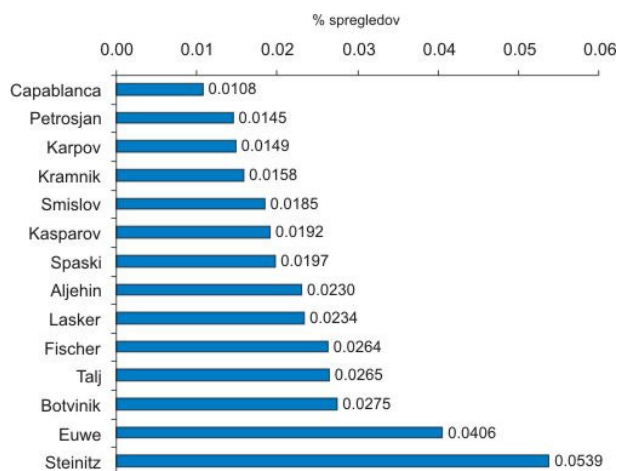


Splošno gledano, rezultati analize delujejo smiselno in se jih s strani šahovskih strokovnjakov da prav lepo interpretirati. Kljub temu bodo nekateri rezultati za mnoge prav gotovo presenetljivi. Zmagovalec po osnovnem kriteriju, kjer smo merili povprečne razlike med odigranimi in najblje ocenjenimi potezami, je tretji svetovni prvak, **Jose Raul Capablanca**.

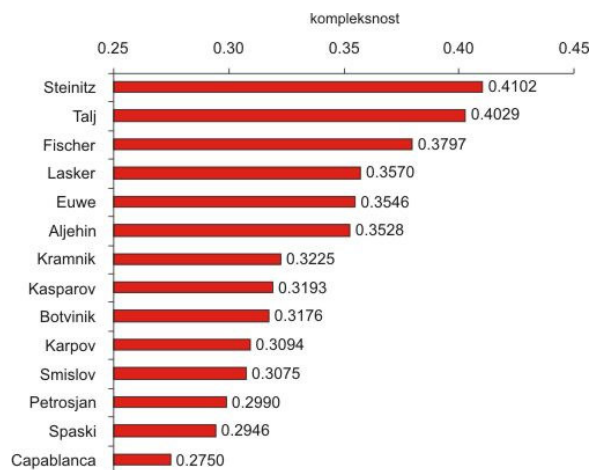
Ta rezultat gre prisoditi relativno nizkim kompleksnostim pozicij, izmerjenih v Capablancinih partijah. Vendar je bil na vrhu tudi pri ostalih lestvicah, kjer smo merili kvaliteto igre in je zaostal le po enem kriteriju (sicer izredno pomembnemu), in sicer pri izračunanih performansah igralcev v primeru, da bi vsi imeli opraviti z enako kompleksnimi pozicijami. Pri slednjem je bil na vrhu sedanji svetovni prvak, **Vladimir Kramnik**. Tako Capablancina kot Kramnik sta izrazito odstopala od ostalih igralcev.



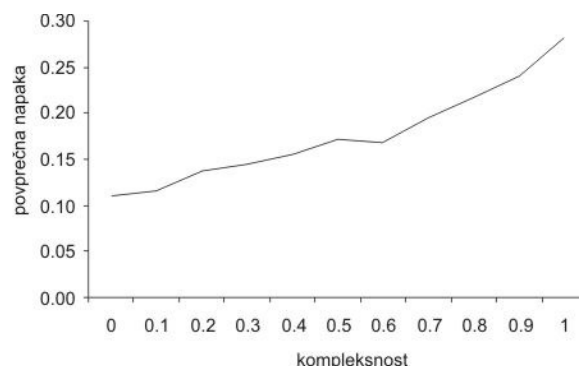
Osnovni kriterij za ocenjevanje svetovnih prvakov je bila **povprečna razlika med odigranimi in najbolj ocenjenimi potezami**. Zmagovalec po tem kriteriju je tretji svetovni prvak, ki je vladal na šahovskem prestolu med leti 1921 in 1927, **Jose Raul Capablanca**. Ta rezultat je po vsej verjetnosti povezan z relativno nizkimi izmerjenimi kompleksnostmi pozicij v Capablancinih partijah, kar se sklada s šahovsko literaturo, povezano z ocenami njegovega stila igre. **Gari Kasparov** v svoji zbirki knjig *Moji veliki predhodniki*, komentirajoč Capablancine partije, celo navaja, da Kubanec občasno sploh ni izgubljal časa z računanjem kompleksnih taktičnih variant, ampak je preprosto dal prednost potezam, ki so pozicijsko bile tako močno utemeljene, da je bilo računanje preprosto odveč. Capablancina opisuje še z naslednjimi besedami: »Zmagoval je vse najpomembnejše partije in dvoboje ter bil neporažen več let zapored (od vseh prvakov je izgubil najmanj partij)« in »njegov stil, eden najbolj kristalno čistih v zgodovini šaha, navdušuje s svojo logiko«.



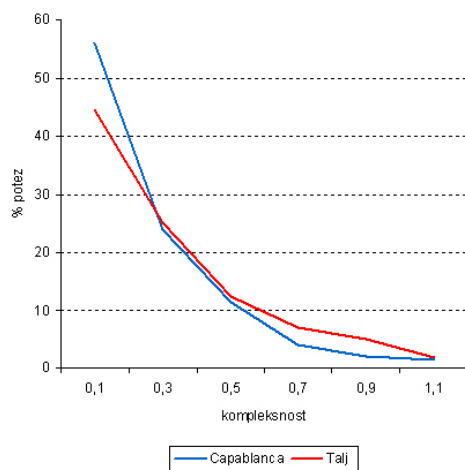
Rezultati **merjenj grobih napak** so podobni. Bilo je pričakovati, da bodo pozicijski igralci v prednosti pred taktičnimi igralci, saj v mirnih pozicijah, s katerimi so se navadno soočali, ni bilo veliko možnosti za spreglede. Opaziti velja odličen rezultat **Tigrana Petrosjana**, ki je splošno znan kot tipičen pozicijski igralec. V skladu z omenjenim je Steinitz, ki je živel v stoletju na oko prijetnega, vendar zato manj korektnega *romantičnega šaha*, zasedel prepričljivo zadnje mesto.



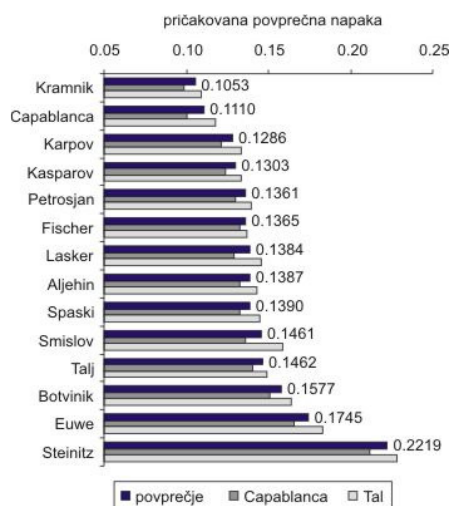
Capablanca je znan po tem, da je igral »preprost« šah in se je izogibal komplikacijam, medtem ko sta se Steinitz in Talj v svojih partijah pogosto soočala z »divjimi« pozicijami. Rezultati **merjenja kompleksnosti** pozicij se povsem skladajo s tako opredeljenim splošnim mnenjem.



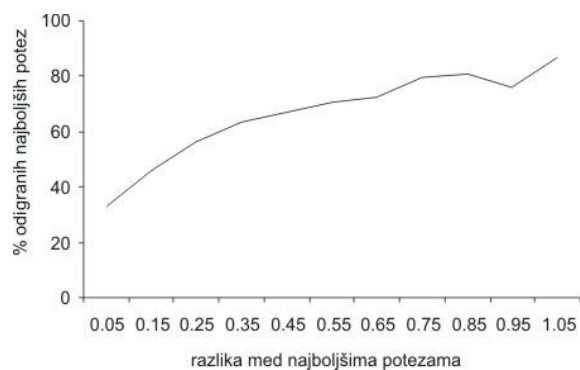
Graf odvisnosti napak od kompleksnosti pozicij jasno kaže na korektnost uporabljene metode za merjenje kompleksnosti pozicij. Igralci so delali manj napak v bolj preprostih pozicijah, povprečne izmerjene napake pa naraščajo s kompleksnostjo pozicij.



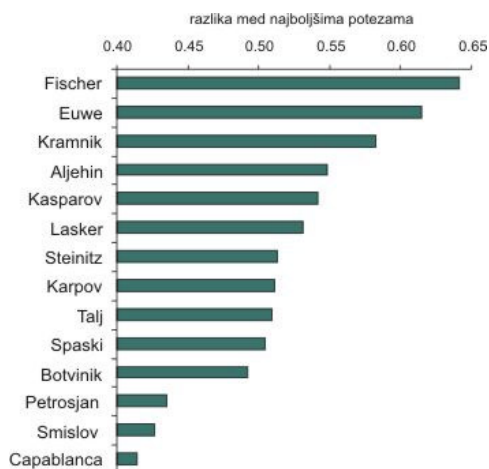
Merilo za ugotavljanje kompleksnosti pozicij smo uporabili, da bi ugotovili porazdelitev potez po posameznih intervalih kompleksnosti, kar je v veliki meri povezano s *stilom igre* igralcev. Capablanca je, kot nazorno kaže zgornja slika, imel veliko manj pogosto opravek z bolj kompleksnimi pozicijami v primerjavi s Taljem, ki je velja za taktičnega igralca.



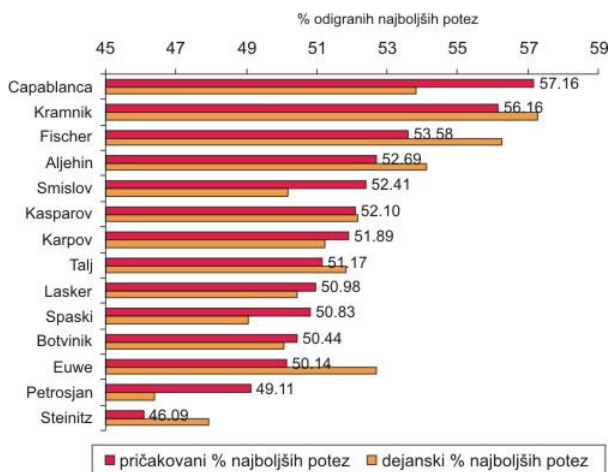
Glavna slabost prejšnjih dveh kriterijev je, kot je že bilo rečeno, da sta nekoliko nepoštena do taktičnih igralcev, ki so v povprečju imeli opravek z bolj kompleksnimi pozicijami, vendar so se v njih tudi bolje znašli od svojih nasprotnikov in so izkoriščali to dejstvo za doseganje odličnih rezultatov. Hoteli smo preveriti, **kako bi se igralci odrezali, če bi vsi imeli opravek z enako kompleksnimi pozicijami**. Zmagovalec po tem kriteriju je bil sedanji prvak, **Vladimir Kramnik**. Kramnik hkrati drži tudi rekord »najboljši rezultat posameznega dvoboja« in sicer je v dvoboju proti Kasparovu (London, 2000) njegova izmerjena povprečna napaka znašala le 0.0903. Prav tako smo skušali ugotoviti, kako bi se igralci odrezali, če bi vsi igrali v stilu Capablanke, Talja itd. Zanimivo je opaziti, da bi Kasparov prehitel Karpova, če bi oba igrala v Taljevem stilu.



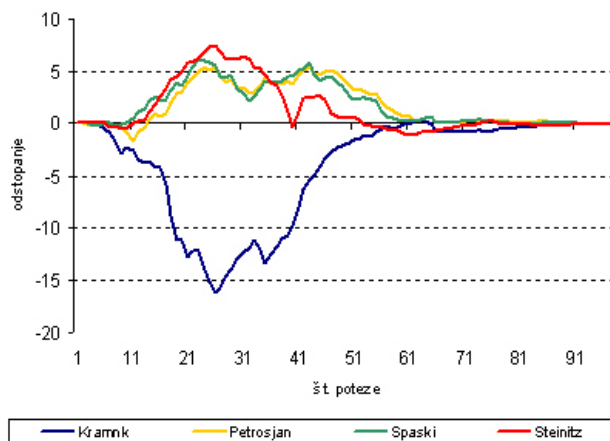
Odstotek najboljših potez sam po sebi ne pove toliko o moči igralca, kot se to morda zdi na prvi pogled. V določenih tipih pozicije je veliko lažje najti najboljšo potezo kot v drugje. Poskusi so pokazali, da je odstotek najboljših odigranih potez tesno koreliran z razliko v ocenah med najboljšima potezama v dani poziciji. Večja kot je razlika, večji je bil izmerjeni uspeh igralcev pri odločanju za najboljšo potezo.



Kramnik, Fischer in Aljehin so največkrat odigrali najboljšo potezo, vendar so v njihovih pozicijah bile tudi relativno visoke razlike med najboljšima potezama. V nasprotju s tem opažanjem je v pozicijah Capablance, ki po odstotku odigranih najboljših potez takoj sledi omenjeni trojici, bila izmerjena v povprečju najmanjša razlika med najboljšima potezama.



Omenjena ugotovljena korelacija je bila podlaga za oblikovanje novega kriterija: **pričakovani odstotek odigranih najboljših potez**, če bi vsi igralci imeli opravka z enakimi razlikami med najboljšima potezama. Ta kriterij predstavlja še drug poizkus spraviti vse prvake na »skupni imenovalac«. Zmagovalec po tem kriteriju je bil znova Capablanca. Skupaj s Kramnikom po rezultatu spet močno izstopata pred ostalimi.



Namen izračunov povprečnega seštevka materiala na šahovnici ni bil povezan z ugotavljanjem kvalitete igre, pač pa smo želeli pridobiti dodatne informacije o stilu igre posameznih igralcev. Predvsem smo želeli odkriti **težnje igralcev po poenostavitvah pozicije**. Med tistimi, ki odstopajo od povprečja, je Kramnik očitno imel največkrat opravka z najmanj materiala na šahovnici. Nasprotno bi lahko trdili za Steinitza, Spaskega in Petrosjana.

Izvirni članek, ki natančno podaja podrobnosti v zvezi z uporabljenimi metodologijo, je bil objavljen v znanstveni reviji *International Computer Games Association (ICGA) Journal* ter bil predstavljen na mednarodni konferenci *Computers and Games* maja letos v Torinu.

Reference

- [1] Matej Guid in Ivan Bratko. *Computer Analysis of World Chess Champions*. *ICGA Journal*, Vol. 29, No. 2, Jun. 2006, str. 65-73.
- [2] Jeff Sonas. *Chessmetrics*. <http://www.chessmetrics.com>.
- [3] Ken Thompson. *Computer chess strength*. *Advances in Computer Chess 3*, Pergamon Press, str. 55-56, 1982.
- [4] Gari Kasparov in Dimitrij Plisecki. *Moji veliki predhodniki*. Šahohlačnik, Maribor.