

UNIVERZA V LJUBLJANI
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

RAČUNALNIŠKA ANALIZA
SVETOVNIH ŠAHOVSKIH PRVAKOV

MATEJ GUID

Delo je opravljeno v skladu s Pravilnikom o podeljevanju Prešernovih nagrad študentom, pod mentorstvom prof. dr. Ivana Bratka.

Ljubljana, september 2005

Povzetek

Ocene, kdo je bil najboljši svetovni šahovski prvak vseh časov, temeljijo predvsem na analizah njihovih partij s strani šahovskih vele mojstrov, le-te pa so pogosto *subjektivne narave*.

Pojavitev vse močnejših šahovskih programov, ki se dandanes lahko že kosajo z najboljšimi šahisti na svetu, v nekaterih elementih igre pa jih celo že presegajo, omogoča dobiti *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.

Pri ocenjevanju svetovnih šahovskih prvakov sem imel drugačen pristop: zanimala me je *kvaliteta njihove igre*, ki sem jo ocenjeval s pomočjo računalniške analize posameznih *potez*. V ta namen sem uporabil računalniški šahovski program Crafty, ki slovi kot najmočnejši šahovski program odprte kode na svetu. Program sem ustrezno spremenil tako, da ga pri ocenjevanju posameznih potez nisem časovno omejeval, pač pa sem mu določil *maksimalno fiksno globino* preiskovanja iskalnega drevesa. S tem sem omogočil njegovo izvajanje na različno hitrih računalnikih, ne da bi to vplivalo na rezultate, hkrati pa je tako samodejno namenjal več časa zahtevnejšim pozicijam.

Svetovne prvake sem med seboj primerjal po različnih kriterijih, kot so odstopanja odigranih potez od izbranih računalnikovih potez, ugotavljanje napak in spregledov ter odstotek odigranih najboljših potez. Poleg ocenjevanja odigranih potez sem ugotavljal tudi težavnost pozicij, s katerimi so se igralci soočali, in odstopanja med najboljšima predlaganima potezama v njih. Na osnovi teh ocen sem ugotavljal pričakovano kvaliteto igre svetovnih prvakov pri uravnoveženih pogojih, kar predstavlja poskus, da bi prvake kljub njihovim različnim stilom igre pri ocenjevanju "spravil na skupni imenovalce". Rezultati analize so mi omogočili tudi pregled vzponov in padcev kvalitete igre v karierah igralcev in ugotavljanje njihove forme v posameznih dvobojih.

Verodostojnost Craftyja kot ocenjevalca, kot tudi korektnost uporabljenih metod dela, sem dokazal z ugotovitvijo sovpadanja napak igralcev z dejanskimi rezultati partij ter s potrditvijo odvisnosti napak od kompleksnosti pozicij.

Kdo je torej bil najboljši svetovni prvak vseh časov? Zmagovalec po glavnem kriteriju, pri katerem smo merili povprečna odstopanja med ocenami odigranih potez in najboljše ocenjenih potez s strani računalnika, je Jose Raul Capablanca, ki je bil svetovni prvak med letoma 1921 in 1927. Na vrhu je bil tudi pri vseh ostalih kriterijih, s katerimi smo merili kvaliteto igre, le pri ugotavljanju kvalitete igre pri različnih stilih igranja ga je prehitel najmlajši svetovni prvak Vladimir Kramnik. Oba sta tudi sicer izrazito odstopala od ostalih.

Ključne besede: računalniški šah, šahovski programi, šah, svetovni šahovski prvaki

Abstract

Estimates about who was the strongest World Chess Champion of all times, are based primarily on the analyses of their games done by chess grandmasters and these are often subjective.

The emergence of better and better chess programs, which can nowadays already cope with the best chess players in the world and even surpass them in certain elements, help gaining more objective answers to this question. Regardless of that, in past researches computers were mainly used for processing statistical data of the chess games' results.

I had a different approach at evaluating World Chess Champions; I was mainly interested in the quality of their game, which I was evaluating with the help of computer analysis of individual moves. For this purpose I used a computer chess program, Crafty, which is renowned for being the most powerful open source chess program in the world. I altered the program so that it would not have a time limit at evaluating individual moves and set a maximum fixed depth of processing the search tree. By doing so, I enabled it to function on any computer regardless of its speed, without that having to influence the results, while at the same time forcing it to dedicate more time to more difficult positions.

I compared the World Champions based on different criteria, such as deviations of played moves versus the moves chosen by a computer, estimating errors and blunders, and the percentage of played out best moves. In addition to evaluating played moves, I also estimated the difficulty of positions with which the players were faced, and deviations of best two suggested moves in them. Based on these estimates I was ascertaining the expected game quality of World Champions under balanced conditions, which represents an attempt to bring the champions to a common denominator while evaluating regardless of their different game styles. The results of the analyses also provided me with an overview of ups and downs of game quality in the players' careers and ascertaining their form in individual duels.

I proved Crafty's credibility as an evaluator as well as the correctness of the used working methods, by finding that the players' errors coincided with actual game results and by confirming the dependence of errors on the complexity of the positions.

So who was the best World Champion of all times? The winner according to the main criterion, where we measured average deviations between evaluations of played moves and best evaluated moves according to the computer, is Jose Raul Capablanca, who was the World champion during 1921 and 1927. He was also on top according to all other criteria, where we measured game quality, and was only beaten in one criterion, game quality in different game styles, by the youngest World Champion, Vladimir Kramnik. Both of them distinctly deviated from the rest.

Keywords: computer chess, chess programs, chess, World Chess Champions

Kazalo

POVZETEK	2
POVZETEK V ANGLEŠKEM JEZIKU	3
UVOD.....	5
1.1 KDO JE BIL NAJBOLJŠI SVETOVNI ŠAHOVSKI PRVAK VSEH ČASOV?.....	5
1.2 NA KRATKO O DELOVANJU ŠAHOVSKIH PROGRAMOV.....	6
1.2.1 Iskalna drevesa in minimaks.....	6
1.2.2 Rezanje dreves in alfa-beta iskalni algoritem.....	8
1.2.3 Tabela transpozicij.....	9
1.2.4 Ničelna poteza.....	10
1.2.5 Poglobljevanje iskanja.....	10
1.3 ŠAHOVSKI PROGRAM CRAFTY.....	12
METODE DELA.....	14
2.1 SVETOVNI ŠAHOVSKI PRVAKI.....	14
2.2 OBRAVNAVANE PARTIJE.....	15
2.3 DOSEDANJE RAZISKAVE.....	16
2.4 NAŠ PRISTOP.....	18
2.5 SPREMENJENI CRAFTY.....	18
2.6 OCENJEVANJE POTEZ.....	20
2.7 ZAJETI PODATKI.....	21
2.8 PODATKOVNA BAZA.....	22
2.9 KRITERIJI OCENJEVANJA.....	23
2.9.1 Osnovni kriteriji.....	25
2.9.2 Kombinirani kriteriji.....	34
2.10 VERODOSTOJNOST CRAFTYJA KOT OCENJEVALCA.....	38
2.10.1 Crafty kot ocenjevalec napak.....	38
2.10.2 Crafty kot ocenjevalec kompleksnosti pozicij.....	44
2.11 METODE DELA S TEHNIČNEGA VIDIKA.....	45
PREDSTAVITEV REZULTATOV	47
3.1 NAPAKE IGRALCEV.....	47
3.1.1 Povprečna napaka pri vsaki potezi.....	47
3.1.2 Primerjava igralcev po dvobojih.....	49
3.1.3 Primerjava povprečnih napak pri partijah posameznih dvobojev.....	50
3.1.4 Število grobih napak (spregledov).....	51
3.2 TEŽAVNOST ODLOČANJA (KOMPLEKSNOŠT POZICIJE).....	52
3.3 ODSOTEK ODIGRANIH NAJBOLJŠIH POTEZ.....	55
3.4 NAPAKA IGRALCEV PRI URAVNOTEŽENI KOMPLEKSNOŠTI POZICIJ.....	58
3.5 MATERIAL TEKOM PARTIJE.....	60
SKLEP	63
ZAHVALA.....	65
LITERATURA.....	66
DODATEK	67
A IZVORNA KODA SPREMENJENEGA CRAFTYJA.....	67
B ANALIZA NAPAK SVETOVNIH PRVAKOV PO POSAMEZNIH DVOBOJIH.....	70

1

Uvod

1.1 Kdo je bil najboljši svetovni šahovski prvak vseh časov?

Vprašanje iz naslova že dolgo časa buri duhove svetovne šahovske javnosti. Navadno ima vsakdo izmed šahistov glede tega svoje mnenje, ki temelji predvsem na preanaliziranih partijah najboljših šahistov v različnih obdobjih zgodovine. Vendar so človeške analize šahovskih partij pogosto zelo *subjektivne narave*. Trinajsti svetovni šahovski prvak Gari Kasparov v svoji še nedokončani zbirki knjig z naslovom *Moji veliki predhodniki (My great predecessors)* [3] podrobno analizira številne partije svetovnih prvakov in bo na koncu prav gotovo podal svojo oceno, kdo med njimi je bil najboljši, vendar bo le-ta sprejeta predvsem kot njegovo osebno *mnenje*, pa čeprav zelo cenjeno v šahovskem svetu.

Vse močnejši računalniki in prihod računalniških šahovskih programov, ki se dandanes lahko že enakovredno kosajo s šahovskimi vele mojstri, omogočajo pridobitev objektivnejših ocen glede zastavljenega vprašanja. Toda večina dosedanjih tovrstnih raziskav je bila usmerjena predvsem v statistične obdelave rezultatov partij svetovnih prvakov. Le-ta pa ima številne pomanjkljivosti, saj rezultati partij marsikdaj ne izražajo najboljše dejanske kvalitete igralca, prav tako pa je skoraj nemogoče oceniti moč njihovih nasprotnikov, še posebej tistih, ki so igrali šah še pred uvedbo ratinškega sistema točkovanja.

Naš pristop je bil drugačen: svetovne prvake smo ocenjevali po kvaliteti njihove igre in sicer s pomočjo analize posameznih *potez*. V ta namen sem uporabil računalniški šahovski program Crafty [6, 8], ki slovi kot najmočnejši šahovski program odprte kode na svetu. Program sem ustrezno spremenil tako, da ga je bilo mogoče poganjati na različno hitrih računalnikih, ne da bi to vplivalo na rezultate. Tako je bilo mogoče izvesti več analiz in višje kvalitete. Hkrati pa

je spremenjeni program samodejno namenjal več časa pozicijam, ki so zahtevnejše za ocenjevanje.

Za ocenjevanje kvalitete igre svetovnih prvakov sem razvil različne metode dela in določil več kriterijev. Njihov namen ni bil samo ugotoviti napak v igri prvakov, pač pa tudi dognati, s kako kompleksnimi pozicijami so se navadno soočali, in tako preveriti težavnost njihovega odločanja, oceniti njihove težnje po menjavah figur itd. Na koncu sem s pomočjo simulacij ugotavljal, kolikšne napake bi igralci delali pri različnih stilih igranja in kolikšen odstotek najboljših potez bi odigrali, če bi vsi imeli opravka z enakimi pozicijami. Analiziral sem tudi vzpone in padce v kvaliteti igre v njihovih karierah ter pri nekaterih zanimivih dvobojih tudi skušal opredeliti nihanja v njihovi formi od začetka do konca dvoboja.

Preveril sem tudi sovpadanje ugotovljenih napak z dejanskimi rezultati partij in odvisnost napak od kompleksnosti pozicij. Na ta način sem uspel dokazati tako korektnost uporabljenih metod kot tudi verodostojnost Craftyja kot ocenjevalca. Slednje je bilo še posebej pereče vprašanje pred pričetkom dela, saj Crafty po šahovski moči velja za slabšega od vsaj nekaterih svetovnih prvakov, vendar se je izkazalo, da je, med drugim tudi zaradi velikega števila preanaliziranih potez, Crafty dovolj zanesljiv v svojih ocenah za verodostojnost rezultatov.

1.2 Na kratko o delovanju šahovskih programov

Namen tega poglavja ni podroben opis delovanja računalniških šahovskih programov, pač pa predstavitev nekaterih konceptov, razumevanje katerih je bilo povezano z metodami dela, ki sem jih uporabil za doseg zastavljenih ciljev.

1.2.1 Iskalna drevesa in minimaks

Pri šahu si tako kot pri večini tipičnih iger, kjer se soočita dva nasprotnika, računalniški programi pri izbiri najboljših potez pomagajo z izgradnjo iskalnih dreves. *Koren* tega drevesa pri šahu predstavlja trenutno pozicijo na šahovnici, *vozlišča* so šahovske pozicije, poteze so

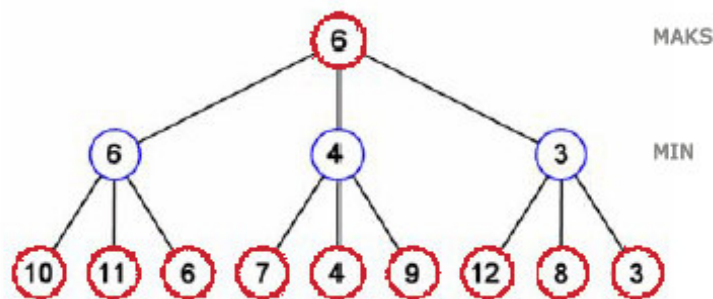
predstavljene kot *povezave* med vozlišči, *listi* pa predstavljajo končne pozicije izgrajenega drevesa.

Drevo, ki bi obsegalo vse možne pozicije v šahovski partiji, od prve poteze pa vse do končnih pozicij, bi vsebovalo približno 10^{120} vozlišč in listov [2], kar je več od števila atomov v znanem vesolju! Izgradnja tako velikega drevesa seveda ni mogoča, zato programi izgradijo iskalna drevesa le do določene globine. Listi v teh drevesih tako navadno niso končne pozicije šahovskih partij, pač pa pozicije pri *maksimalni globini* preiskovanja.

S *preiskovanjem* iskalnih dreves programi skušajo ugotoviti vrednost izhodiščne pozicije in določiti najboljšo potezo. Preiskovanje poteka tako, da računalnik dodeljuje ocene posameznim listom in vozliščem izgrajenih dreves, pri čemer uporablja *najprej v globino* (*depth-first*) strategijo. Pri ocenjevanju uporablja *ocenjevalno funkcijo* (*evaluation function*), vendar z njo ocenjuje le vrednosti listov, medtem ko ocene vozlišč računa iz ocen njihovih neposrednih *naslednikov*, torej listov in vozlišč, do katerih vodijo njihove povezave.

Pri tem se programi držijo *minimaksa* (*mini-max*), ki izhaja iz dejstva, da kar je dobro za enega igralca, mora biti slabo za drugega. Označimo npr. koren iskalnega drevesa kot MAKS, vsa vozlišča, do katerih vodijo povezave iz korena, kot MIN, njihove naslednike spet kot MAKS itd. Dodeljevanje ocen vozliščem poteka na naslednji način:

- vozliščem, ki so označeni z MIN, dodelimo *najmanjšo oceno* njihovih naslednikov,
- vozliščem, ki so označeni z MAKS, dodelimo *največjo oceno* njihovih naslednikov.



Slika 1.1: Iskalno drevo pri algoritmu minimaks

Na tak način se od spodaj navzgor ocenijo vrednosti vseh vozlišč v poddrevesih, ki izhajajo neposredno iz korena drevesa, nakar se izračuna še vrednost v korenu. Le-ta predstavlja

končno oceno trenutne pozicije, povezava, ki vodi do vozlišča, katerega oceno smo dodelili korenu, pa je izbrana najboljša poteza (več v [1, 2]).

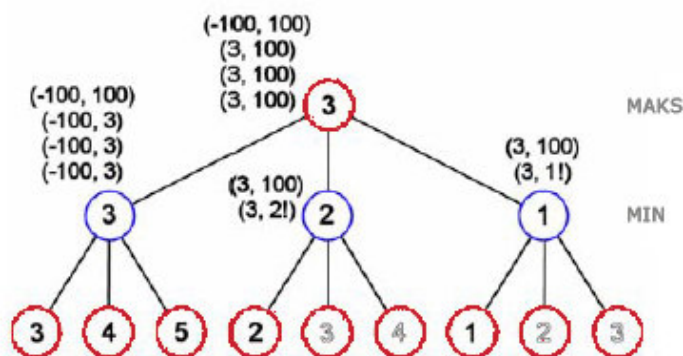
1.2.2 Rezanje dreves in alfa-beta iskalni algoritem

Čeprav iskalna drevesa predstavljajo le vrh celotnega *drevesa igre (game tree)*, so pri igrah kot je šah še vedno ogromna. V običajni šahovski poziciji je v povprečju na voljo okrog 30 potez. To pomeni, da bi iskalno drevo, ki bi obsegalo vse pozicije, do katerih bi bilo mogoče priti pri iskanju do le deset *polpotez* globoko, vsebovalo več kot 500 bilijonov listov in vozlišč. Če privzamemo, da bi računalnik bil zmožen oceniti 500.000 pozicij na sekundo (kar je za današnje razmere veliko število), bi odločitev za vsako posamezno potezo trajala več kot milijardo sekund!

Rešitev predstavlja *rezanje dreves (tree pruning)* oz. nepregledovanje delov iskalnega drevesa, ki za oceno pozicije niso pomembni. Najbolj uporabljan algoritem za ta namen pri šahovskih programih predstavlja *alfa-beta (alpha-beta)* algoritem. Le-ta še vedno upošteva princip minimaks, temelji pa na načelu, da za poteze, ki so *dovolj dobre*, ni potrebno ugotoviti točno kako dobre so in obratno, za poteze, pri katerih vemo, da so *slabše* od že ocenjenih, ne potrebujemo vedeti njihove natančne ocene (več v [1, 2]).

Tako pri označitvi vozlišč z MIN in MAKS na način, kot smo to storili prej, velja [2]:

- *alfa* vozlišča MAKS je spodnja meja, določena z vrednostjo tega vozlišča,
- *beta* vozlišča MIN je zgornja meja, določena z vrednostjo tega vozlišča,



Slika 1.2: Delovanje algoritma alfa-beta, s klicajem so označeni rezi poddreves

Algoritem lahko opišemo z naslednjimi pravili:

- iskanje poddrevesa MIN pod vozliščem MAKS se preneha, če je $\alpha_{\text{MAKS}} \geq \beta_{\text{MIN}}$,
- iskanje poddrevesa MAKS pod vozliščem MIN se preneha, če je $\alpha_{\text{MAKS}} \geq \beta_{\text{MIN}}$,
- ažurira se alfa ali beta vozlišča, katerega nasledniki so bili pregledani.

Začetni vrednosti alfa in beta v korenu drevesa sta pri običajnem alfa-beta iskanju $-\infty$ in ∞ . Poskus izboljšave algoritma alfa-beta predstavljajo *aspiracijska okna* (*aspiration search*), kjer sta ti dve vrednosti čim boljše približki idealnih vrednosti alfe in bete za uspešno rezanje. Tako je pri iskanju navadno potrebno pregledati še manj pozicij. Majhna slabost tako popravljenega algoritma je, da je v primeru slabo določenih začetnih vrednosti alfe in bete iskanje potrebno ponoviti. Vendar kljub temu navadno daje boljše rezultate kot osnovni algoritem alfa-beta [5].

Alfa-beta algoritem je tem uspešnejši, čim prej pridejo na vrsto boljše poteze. Na ta način se namreč porežejo večja poddrevesa. V najboljšem primeru, ko bi na prvem mestu vedno bila najboljša poteza, bi program namesto N pozicij moral oceniti le \sqrt{N} pozicij, kar pomeni, da bi pri enakem porabljenem času iskal dvakrat globlje [1, 10].

Za razvrščanje potez lahko uporabimo *iterativno poglobljanje* (*iterative deepening*), kjer se rezultati ene iteracije uporabljajo za razvrščanje naslednje. Rezultati iterativnega poglobljanja se lahko uporabljajo tudi za določanje natančnejših vrednosti pri metodi aspiracijska okna [9].

Še ena metoda za čim boljše razvrščanje je *ubijalska hevrstika* (*killer heuristic*), ki izhaja iz ideje, da je poteza, ki se je pri neki globini izkazala za dobro v enem poddrevesu, pri isti globini pogosto dobra tudi v drugih poddrevesih. *Zgodovinska hevrstika* (*history heuristic*) pa je posplošitev te metode, za vse uspešne poteze pri opravljenem iskanju namreč hrani njihove ocene in jih uporablja pri razvrščanju potez [4, 5, 9].

1.2.3 Tabela transpozicij

Pogosto se zgodi, da zamenjava vrstnega reda več potez vodi do iste pozicije. Če si po prvem iskanju zapomnimo oceno te pozicije, nam naslednjič ni treba preiskovati drevesa pod njo,

ampak zgolj preberemo oceno iz tabele. Ta tabela se imenuje *transpozicijska tabela* (*transposition table*) in je navadno implementirana kot zgoščena tabela [9].

1.2.4 Ničelna poteza

Ničelna poteza (*null move*) je metoda, pri kateri prepustimo še drugo zaporedno potezo nasprotniku in nato preiskujemo drevo pri zmanjšani globini. Ideja te metode je preskočiti pregledovanje pozicij, ki so *dovolj dobre* in tako pospešiti iskanje. Če je rezultat večji od vrednosti bete, opustimo nadaljnje iskanje. V nasprotnem primeru nadaljujemo z iskanjem po normalni poti.

Ničelna poteza vsebuje tudi nevarnost, ki izhaja iz dejstva, da je včasih v šahu ugodno prepustiti potezo nasprotniku. To se v šahovskem žargonu imenuje *nujnica* (*zugzwang*). Ker takšne situacije nastopijo večinoma v končnicah, programi navadno izklopijo to metodo pri prehodu v zaključni stadij partije [4, 5].

1.2.5 Poglobljanje iskanja

Posebno nevarnost za korektnost ocen predstavlja *učinek obzorja* (*horizon effect*). Včasih se nevarne *pretnje* lahko skrivajo malo dlje od največje globine iskanja. Nasprotnik npr. ima lahko, če bi prišel na potezo, na voljo kombinacijo, ki se konča z matiranjem našega kralja. Človek pogosto posebej obravnava pretnje nasprotnika in v primeru zaznave tovrstne pretnje tudi poskrbi za preprečitev njene izvršitve. Pri računalniku pa obstaja nevarnost, da v poziciji, ko bi idejo nasprotnika še lahko preprečil, najde serijo *forsiranih* potez, npr. dajanje šahov, ki se lahko eventualno tudi konča z dobitkom materiala (npr. osvojitvijo kraljice), vendar se na koncu zgodi, da nasprotnikove matne kombinacije ne more več preprečiti. Ta serija potez namreč potisne nasprotnikovo pretnjo preko *meje obzorja* oz. naprej od največje globine iskanja.

Ta problem se rešuje s pomočjo *selektivnih razširitev* (*selective extensions*), ki pomenijo dodatno poglobljeno iskanje oz. nadaljnje razvijanje vozlišč, ki je še posebej smiselno v primerih, ko sumimo v oceno ocenjevalne funkcije v danem vozlišču [4, 5]. Nesmiselno bi npr. bilo zaključiti, da imamo majhno prednost v poziciji na osnovi nekega slabega polja v nasprotnikovem taboru, ko pa ne moremo preprečiti izgubo kraljice čez nekaj potez.

Selektivne razširitve se delijo na:

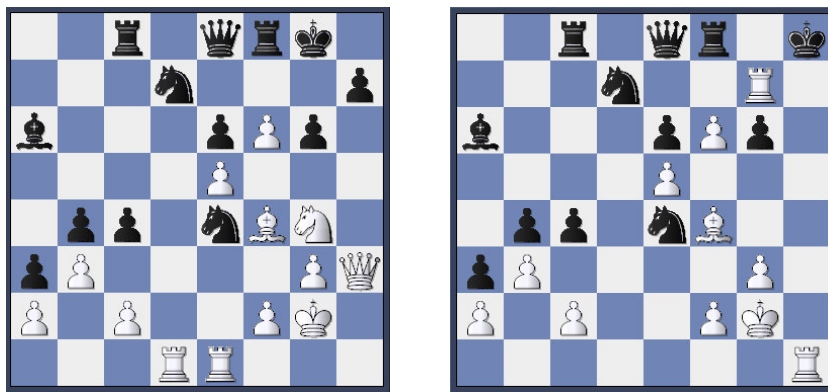
- *iskanje mirovanja* (*quiescence search*),
- *izjemne razširitve* (*singular extensions*).

1.2.5.1 Iskanje mirovanja

Iskanje mirovanja (*quiescence search*) je zelo učinkovita metoda za odpravljanje negativnih posledic učinka obzorja. Ideja je v tem, da z dodatnim poglobljenim iskanjem dosežemo, da se ocenjujejo samo stabilne pozicije. To so pozicije, v katerih ni nobenih direktnih pretenj, kot so npr. poteze, ki vodijo do izgube katere od figur, serije šahov, ki se lahko končajo z matiranjem kralja, promocije kmetov itd. [5]

V ta namen pri doseženi maksimalni globini iskanja preiščemo še vse *forsirane variante*, kot so jemanja, dajanje šahov in umiki pred njimi in pretnje promocij kmetov v katero od figur, dokler se ne pojavi stabilna pozicija, ki jo nato ocenimo.

Šahovski poziciji, prikazani na diagramih na naslednji strani, nazorno prikazujeta učinkovitost iskanja mirovanja. Računalniški programi v poziciji na diagramu levo izredno hitro najdejo serijo potez, ki vodi do mata, čeprav je le-ta oddaljen kar 15 polpotez! Če bi se pri odločanju opirali le na statične ocene pozicije, bi prav gotovo zavrnili zmagovito serijo forsiranih potez, saj ima beli v vseh pozicijah na poti do mata célo damo manj.



Slika 1.3: Demonstracija učinka iskanja mirovanja. Platz-Just, 1972: beli, ki je na potezi (diagram levo), lahko s pomočjo žrtve dame in skakača neubranljivo matira nasprotnega kralja (matna slika je prikazana na diagramu desno).

Do mata vodi naslednja varianta:

29.Dxh7+ Kxh7 30.Th1+ Kg8 31.Sh6+ Kh7 32.Sf7+ Kg8 33.Th8+ Kxf7 34.Th7+ Kg8 35.Tg7+ Kh8 36.Th1#.

1.2.5.2 Izjemne razširitve

Pri *izjemnih razširitvah (singular extensions)* se osredotočimo na poglobljeno iskanje v delu drevesa, kjer smo pri posameznem nasledniku zaznali občutno boljšo oceno kot pri ostalih. To naj bi pomenilo, da se tam dogaja nekaj zanimivega, kar naj bi upravičilo dodatno iskanje, saj bi rezultati tega iskanja lahko odločilno vplivali na oceno izhodiščne pozicije [5].

1.3 Šahovski program Crafty

Crafty [6, 8] velja za najmočnejši program odprte kode na svetu. Pri svojem delu sem uporabljal verzijo 19.2, ki velja za dobro preverjeno in zelo stabilno različico programa.

Avtor Craftyja je dr. Robert Hyatt, profesor na univerzi v Alabami (*University of Alabama*) v Birminghamu, ki se s programiranjem šahovskih programov ukvarja že več kot 30 let. Crafty je neposreden naslednik programa *Cray Blitz*, ki je bil svetovni prvak med računalniškimi šahovskimi programi med letoma 1983 in 1989. Na najnovejšem seznamu organizacije SSDF (*Svenska schackdatorföreningen*) [7], ki se ukvarja z medsebojno primerjavo računalniških

šahovskih programov, Crafty v različici 18.2 iz leta 2001 z ratingom 2617 zaseda visoko 14. mesto (pri čemer niso upoštevane različne verzije istih programov), pred njim pa so uvrščeni samo komercialni šahovski programi.

Odprta koda je bila pogoj za uspešno delo, saj je bilo program potrebno ustrezno spremeniti. Spremenil sem ga tako, da pri ocenjevanju posameznih potez ni bil časovno omejen, pač pa je pri vsaki potezi preiskal iskalno drevo do vnaprej določene *maksimalne fiksne globine* (kako sem spremenil program je podrobneje razloženo v 2.5).

O moči Craftyja je težko razpravljati, vsekakor pa na običajnih osebnih računalnikih s procesorsko močjo prek 2 GHz dosega moč šahovskega velemejstra, medtem ko se pri taktičnih zapletih, podobno kot večina najboljših današnjih programov, lahko meri tudi z najboljšimi šahisti na svetu. Za preprogramirani Crafty velja podobno. Upam si celo trditi, da je spremenjeni Crafty pri analizah partij močnejši od originalnega, saj bolje razporeja čas glede na zahtevnost ocenjevanja posameznih pozicij. Pri vsaki poziciji namreč doseže globino iskanja, ki mu je vnaprej določena in v kateri so ocene že zelo stabilne in zanesljive. Izhajal sem namreč iz velike prednosti, da programa ni bilo potrebno prilagoditi za boj z nasprotniki, torej sem si lahko privoščil, da je analiza kompleksnejših pozicij lahko trajala tudi po več ur, če je to bilo potrebno.

2

Metode dela

2.1 Svetovni šahovski prvaki

Obravnaval sem štirinajst svetovnih prvakov, ki so vladali na svetovnem šahovskem prestolu od prvega uradnega dvoboja za svetovnega prvaka leta 1886 pa vse do današnjih dni. Pri tem sem upošteval t.i. *klasično verzijo* svetovnih prvenstev. Po razpadu svetovne šahovske federacije leta 1993 na dva dela sta namreč vedno prisotna po dva svetovna prvaka. Vendar se je pri tej verziji, ki jo priznava večina šahistov po svetu, bolj ohranil tradicionalen pristop k organizaciji dvobojev za svetovnega prvaka. Izzivalca se tako določi po več dvobojih kandidatov, kateremu sledi zaključni dvoboj med njim in aktualnim prvakom. Poleg tega je sedanjí prvak po tej verziji Vladimir Kramnik naslov prevzel direktno od trinajstega svetovnega prvaka Garija Kasparova, potem ko ga je leta 2000 premagal v dvoboju za svetovnega prvaka v Londonu.

Igralec	Država	Obdobje vladavine
Steinitz Wilhelm	Čehoslovaška	1886 - 1894
Lasker Emanuel	Nemčija	1894 - 1921
Capablanca Jose Raul	Kuba	1921 - 1927
Aljehin Aleksander	ZSSR	1927 - 1935, 1937 - 1946
Euwe Maks	Nizozemska	1935 - 1937
Botvnik Mihail	ZSSR	1948 - 1957, 1958 - 1960, 1961 - 1963
Smislov Vasilij	ZSSR	1957 - 1958
Talj Mihail	ZSSR	1960 - 1961
Petrosjan Tigran	ZSSR	1963 - 1969
Spaski Boris	ZSSR	1969 - 1972
Fischer Robert James	ZDA	1972 - 1975
Karpov Anatolij	ZSSR	1975 - 1985
Kasparov Gari	ZSSR	1985 - 2000
Kramnik Vladimir	Rusija	2000 -

Tabela 2.1: Seznam vseh svetovnih prvakov in obdobja njihovih vladavin na svetovnem šahovskem prestolu

2.2 Obravnavane partije

Za vsakega od obravnavanih šahistov je bilo potrebno izbrati njegove partije, ki bi čim boljše predstavile njegov način igre in odrazile njegovo šahovsko moč. V ta namen sem izbral partije njihovih neposrednih dvobojev za naslov šahovskega prvaka, bodisi so bili v njih v vlogi izzivalca bodisi v vlogi branitelja naslova. Ti dvoboji so namreč vedno bili vzeti izredno resno, povezani so bili tako z najvišjim naslovom kot z denarnimi nagradami in navadno terjajo tudi po več mesecev priprav nanje. Kvaliteta partij v teh dvobojih je tako običajno precej višja kot pri ostalih partijah. Iz tega razloga sem tudi obravnaval *samo* partije iz teh dvobojev, saj bi vključevanje partij iz ostalih šahovskih turnirjev lahko bistveno popačilo rezultate posameznih igralcev. Izjemi sta Fischer in Kramnik, ki sta imela precej manj partij od ostalih (Fischer 20, Kramnik 29) in je torej obstajala nevarnost, da število obravnavanih partij pri njiju ne bi dalo najbolj zanesljivih rezultatov. Vendar je potrebno omeniti, da so tudi dodane partije bile skrbno izbrane. Gre za partije iz dvobojev kandidatov za izzivalca svetovnega prvaka, ki so ju neposredno vodile do dvobojev za svetovnega prvaka in ki so po svoji naravi zelo sorodne samim neposrednim dvobojem za najvišji naslov. Število obravnavanih partij vsakega igralca je razvidno iz tabele 2.2.

Igralec	Št. partij
Karpov	197
Kasparov	197
Botvinik	177
Aljehin	145
Steinitz	115
Lasker	112
Euwe	80
Petrosjan	69
Smislov	69
Spaski	67
Capablanca	48
Talj	42
Fischer	41
Kramnik	38

Tabela 2.2: Število obravnavanih partij pri vsakem od igralcev

2.3 Dosedanje raziskave

Dosedanje raziskave so bile v glavnem usmerjene v statistične obdelave rezultatov igralcev v izbranih partijah. Vendar le-te ne izražajo dejanske moči igralcev niti kvalitete igre v teh partijah iz več razlogov, naj navedem dva bistvena:

- igralci so imeli različno močne nasprotnike, kar nedvomno vpliva na rezultat,
- v šahu lahko že ena sama napačna poteza odločilno vpliva na končni rezultat, pa čeprav so lahko ostale poteze odlične, torej lahko do istega rezultata vodijo poteze povsem različnih kvalitete.

Omeniti sicer velja tudi, da so včasih igralci imeli na voljo več časa za razmišljanje, res pa je tudi, da so imeli na razpolago manj kvalitetne literature in slabše možnosti za šahovsko izobraževanje in izpopolnjevanje.

Rezultatska uspešnost igralcev je prikazana v tabeli 2.3.

Igralec	Št. partij	Št. zmag	Št. porazov	Število remijev	Uspeh (%)
Fischer	41	24	3	14	75,61
Lasker	112	52	16	44	66,07
Aljehin	145	44	26	75	56,21
Kasparov	197	31	23	143	52,03
Petrosjan	69	13	11	45	51,45
Botvink	177	46	41	90	51,41
Capablanca	48	7	6	35	51,04
Karpov	197	32	28	137	51,02
Smislov	69	18	17	34	50,72
Kramnik	38	4	4	30	50,00
Steinitz	115	43	43	29	50,00
Talj	42	11	12	19	48,81
Spaski	67	11	15	41	47,01
Euwe	80	16	32	32	40,00

Tabela 2.3: Uspešnost igralcev v obravnavanih partijah

Kot bomo videli kasneje, rezultati ne sovpadajo z dejansko kvaliteto njihove igre, izmerjene z našimi metodami, ki temeljijo na analizi posameznih *potez* igralcev. Te metode so podrobneje opisane v nadaljevanju.

Do sedaj najodmevnejši in najcelovitejši poskus določitve najboljšega šahista vseh časov s pomočjo statistične obdelave rezultatov je nedavno izvedel Jeff Sonas, ki se je v zadnjih letih uveljavil kot številka ena na svetu, ko gre za statistične raziskave v šahu. Sonas je v ta namen ustvaril posebno rating lestvico in pri tem zajel šahovske partije vse od leta 1840 naprej [12]. Rating je obračunan za vsak mesec posebej, formule za njegov izračun pa upoštevajo tudi aktivnost igralca. Igralčev rating tako za razliko od klasičnega FIDE ratinga v primeru prenehanja igranja prične padati.

Igralec	Rating
Fischer	2895
Kasparov	2886
Botvinik	2885
Lasker	2878
Capablanca	2877
Aljehin	2860
Karpov	2848
Kramnik	2826
Steinitz	2826
Smislov	2800
Talj	2799
Petrosjan	2796
Spaski	2773
Euwe	2769

Igralec	Št. mesecev na vrhu
Lasker	292
Kasparov	263
Steinitz	173
Botvinik	131
Aljehin	122
Fischer	109
Karpov	100
Capablanca	85
Smislov	59
Talj	38
Petrosjan	33
Euwe	14
Spaski	6
Kramnik	0

Tabela 2.4 in 2.5: Najvišji doseženi ratingi svetovnih prvakov in število mesecev, ki so jih prebili na vrhu rating lestvice Jeffa Sonasa.

Zgornji tabeli prikazujeta le podatke o svetovnih prvakah, obravnavanih v tem delu, vendar tudi sicer velja, da so na samem vrhu pri raznih lestvicah, pridobljenimi iz Sonasovih statistik (npr. najboljši performansi na turnirjih in v dvobojih, najvišji rating v različnih obdobjih itd.), izključno svetovni prvaki.

Kljub vsemu pa je iz prikazanih rezultatov težko sklepati o kvaliteti igre posameznih igralcev. Steinitz je npr. na podlagi svoje rezultatske uspešnosti svoj najvišji rating dosegel aprila leta 1876, ko je splošna kvaliteta igre v šahovskih partijah močno zaostajala za današnjo, medtem ko je na rezultate v desni tabeli precej vplivalo tudi v katerem obdobju je igralec živel, saj so bili nekateri pogosto v senci najboljšega v njihovem času, drugi pač ne. Hkrati so nekateri bili aktivni več časa, drugi manj.

2.4 Naš pristop

Naš namen je bil dobiti objektivnejše rezultate, kot so jih lahko ponudile raziskave, ki so temeljile zgolj na statističnih obdelavah rezultatov. Ideja je bila v tem, da bi primerjave med igralci izvedli s pomočjo ocenjevanja njihovih odigranih *potez*.

V ta namen sem uporabil program odprte kode Crafty, ki ima vgrajeno možnost dodeljevanja numeričnih vrednosti za posamezne poteze. Prav tako ima na voljo ukaz, ki omogoči, da se ocene posameznih potez beležijo v za ta namen izbrano datoteko.

Vendar pa je imel Crafty določene pomanjkljivosti za naš namen. Tako ga je bilo pri ocenjevanju posameznih potez mogoče le *časovno omejiti* in ne kako drugače. Omejevanje časa za ocenjevanje potez je vsekakor potrebno, saj računalnik pri ocenjevanju neprestano povečuje globino iskanja in bi pozicije ocenjeval v nedogled, če ne bi bil pri tem kakorkoli omejen. Toda časovno omejevanje v smislu določitve fiksnega števila sekund, namenjenih ocenjevanju ene poteze, ima številne pomanjkljivosti, dve najpomembnejši sta:

- računalnik namenja vsem pozicijam enako časa, tako preprostim kot kompleksnejšim,
- program bi na različno hitrih računalnikih dal različne rezultate.

Slednja pomanjkljivost je prav gotovo razlog, da se tovrstnih raziskav do sedaj niso lotevali, saj bi analiza npr. tisoč partij na enem samem računalniku, pa čeprav hitrem za današnje standarde, trajala več mesecev.

Craftyja je torej bilo potrebno spremeniti tako, da je ustrezal našim potrebam. Kako sem to storil, je podrobneje opisano v naslednjem poglavju.

2.5 Spremenjeni Crafty

Program sem iz razlogov, ki so bili podrobneje opisani v prejšnjem poglavju, preprogramiral tako, da ga pri ocenjevanju nisem časovno omejeval, pač pa sem mu določil *maksimalno*

fiksno globino iskanja za vsako potezo. Na ta način računalnik ne preneha ocenjevati vrednosti posamezne poteze, dokler ne pregleda drevesa preiskovanja do globine, ki mu jo določimo. Tako dosežemo naslednje:

- računalnik v *kompleksnejših pozicijah*, kjer je potrebno obdelati iskalna drevesa večjih razsežnosti za pridobitev točnejše ocene, samodejno nameni več časa,
- program lahko poganjamo na *različno hitrih računalnikih* in obenem pri istih ocenjevanih pozicijah dobimo enake rezultate; na počasnejših računalnikih pride do rezultatov počasneje kot na hitrejših računalnikih, vendar so rezultati povsod enaki.

Izbrana je bila maksimalna fiksna globina iskanja 12 polpotez. Empirično se je pokazalo, da se ocene od te globine dalje zelo malo spreminjajo in bi torej povečanje globine preiskovanja le minimalno pripomoglo h kvaliteti pridobljenih ocen, pri čemer bi se čas ocenjevanja zaradi eksponentne narave preiskovalnih dreves občutno povečal. Večje globine iskanja si zaradi omejenega časa, ki sem ga imel na voljo, tako tudi ne bi mogel privoščiti.

Maksimalno fiksno globino iskanja ob prehodu v *končnico* povečamo na 13 polpotez. Pri tem sem upošteval Craftyjevo definicijo končnice: partija preide v končnico, ko je skupna numerična vrednost tako belih kot črnih figur na šahovnici (brez kmetov) manjša od 15. Pri tem so upoštewane naslednje vrednosti figur: dama 9, trdnjava 5, lovec 3 in skakač 3. Na ta način sem želel doseči še nekoliko točnejše ocene v fazi končnice, za katere računalniški šahovski programi na splošno veljajo za nekoliko slabše v primerjavi s *središčnico*. Pri tem se časi ocenjevanja niso bistveno spremenili, saj so pri ocenjevanju pozicij v tej fazi igre preiskovalna drevesa navadno že občutno manjša.

Omejitev iskanja do vnaprej določene fiksne globine nikakor ne pomeni, da računalnik ne »vidi« dlje od 12 (oz. 13 v fazi končnice) polpotez od ocenjevane pozicije dalje. *Problem horizonta* oz. nevarnost, da bi računalnik spregledal kakšno za oceno pozicije ključno potezo, ki bi nastopila izven okvira preiskanega drevesa, program rešuje s pomočjo *iskanja mirovanja*, ki se samodejno vklopi ob dosegu maksimalne fiksne globine iskanja. Pri tem iskanju program obravnava vse t.i. *forsirane variante*, to so serije potez, ki vključujejo vsa jemanja in dajanja šahov ter odgovore na tovrstne poteze. Rezultati tega iskanja se neposredno upoštevajo pri ocenah ocenjevanih pozicij.

Program je bil spremenjen tudi tako, da podaja izpise pri različnih globinah preiskovanja, kar mi je kasneje omogočilo ugotavljanje težavnosti odločanja v posameznih pozicijah, kot bo opisano v nadaljevanju. S pomočjo spremenjenega obravnavanja pomena varnosti kralja sem tudi dosegel, da je Crafty, ki je po privzeti nastavitvi bolj defenziven, postal v tem pogledu nevtralen oz. ga ne bi mogli opredeliti niti kot napadalnega niti kot obrambnega igralca oz. ocenjevalca. Prav tako sem poskrbel še za večjo informativnost izpisov, dodal sem npr. izpis materialnega stanja pri vsakem od nasprotnikov pri vsaki potezi.

Ocenjevanje je potekalo na 36 računalnikih, z vgrajenimi procesorji Intel Celeron s hitrostjo procesiranja 2.8 GHz in 512 MB spomina ter Intel Pentium IV s hitrostjo 2.4 GHz in prav tako 512MB spomina, opravljenih pa je bilo 1397 analiz izbranih partij.

2.6 Ocenjevanje potez

Ocenjevanje vsake partije se je pričelo pri 12. potezi. Današnji računalniški programi namreč slabo ocenjujejo poteze v začetni fazi igre in pri igranju otvoritev uporabljajo *otvoritvene knjižice*. Le-te so različne pri vsakem od programov in temeljijo predvsem na statističnih izračunih uspešnosti v njih vsebovanih potez, odigranih v partijah, ki so bile zajete pri njihovi izgradnji, ne vsebujejo pa izračunanih točnih ocen pozicij. Za dosego našega namena nisem uporabil kakršnekoli otvoritvene knjižnice predvsem iz naslednjega razloga:

- v primeru uporabe otvoritvene knjižnice bi program še vedno računal ocene potez, ki se v njej ne bi nahajale (in kot sem omenil, so izračuni vrednosti potez v začetni faze igre navadno netočni), medtem ko bi za ostale poteze zgolj opredelil, da se nahajajo v otvoritveni knjižnici ter kvečjemu še podal njihovo statistično vrednost.

Partij pa nisem ocenjeval od prve poteze ne samo zato, ker bi pri začetnih potezah na ta način dobil netočne ocene (kar je samo po sebi zadosten razlog za tako odločitev), marveč tudi zato, ker bi igralci mlajših generacij zaradi vseskozi napredujoče *teorije otvoritev* bili v prednosti pred igralci starejših generacij. Začetna faza igre je namreč vse bolj raziskana, napisane so številne knjige različnih specialistov za izbrane otvoritve, ki temeljijo na njihovih izkušnjah in analizah ter še posebej na dolgoletni praksi odigranih partij, ki so se pričele z izbrano

otvoritvijo. Šahisti se prek različnih publikacij tudi neprestano obveščajo o vseh novostih, ki se pojavijo na področjih posamezne otvoritve. Dandanes je napredek na tem področju še veliko večji zaradi uporabe računalniških programov za učinkovito delo z ogromnimi bazami šahovskih partij, pa tudi obveščanje poteka veliko hitreje z uporabo interneta. Današnji velemejstri imajo tako pogosto vnaprej pripravljenih tudi do 20 začetnih potez in včasih celo več!

Pojavilo se je torej vprašanje, pri kateri potezi pričeti z ocenjevanjem. Zgodnejši pričetek ocenjevanja partij bi bil po vsej verjetnosti v prid svetovnim prvakom mlajših generacij, ki so v fazi otvoritve zaradi omenjenih razlogov bolje pripravljeni od njihovih predhodnikov. Vendar pa bi s prepoznim pričetkom ocenjevanja izgubili marsikatero informacijo o kvaliteti igre obojih, saj težavnost sprejemanja odločitev pri izbiranju potez navadno nastopi že pred 15. potezo. Čeprav so imeli svetovni prvaki mlajših generacij navadno prednost v teoretični podkovanosti v fazi otvoritve, pa tudi ne gre prezreti dejstva, da so njihovi predhodniki imeli na voljo več časa za razmišljanje, kot sem že omenil.

Naš namen je bil predvsem ugotavljati čim objektivnejšo vrednost povlečenih potez, zato sem pri tej odločitvi izhajal predvsem iz ugotovitev, pri kateri potezi se lahko na Craftyja že lahko zanesemo kot na dovolj verodostojnega ocenjevalca. S pomočjo primerjav ocen z ostalimi računalniškimi šahovskimi programi ter z ocenami pozicij s strani človeških velemejstrov se je pokazalo, da je 12. poteza primerna izbira, hkrati pa predstavlja dober kompromis v povezavi z navedenimi dilemami.

2.7 Zajeti podatki

Rezultati analiz so se beležili neposredno v izbrane datoteke v *html formatu*. Poskrbel sem za izpis *diagramov* oz. slik pozicij po vsaki potezi, kar mi je omogočilo sprotni nadzor nad korektnostjo dobljenih rezultatov

S pomočjo preprogramiranega Craftyja sem pri vsaki potezi ocenjevanega igralca zajel naslednje podatke:

- Za vsako od globin iskanja od 2 do 12 (oz. 13, ko je partija v fazi končnice):
 - najboljše ocenjena poteza in njena ocena,
 - druga najboljše ocenjena poteza in njena ocena,
 - odigrana poteza in njena ocena.

Pri vsaki potezi je navedena tudi *kritična varianta* oz. serija izračunanih najboljših potez za obe strani, ki vodi do dodeljene ocene. V primeru, da je odigrana poteza enaka kateri od zgornjih dveh, pri izpisu ne pride do podvajanja.

- Izpis materialnega stanja pri obeh igralcih (od prve poteze naprej).

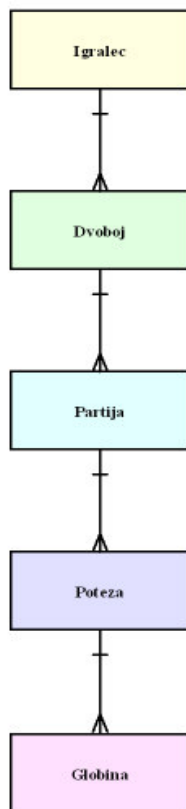
O vsaki partiji pa sem imel na voljo naslednje podatke:

- naslov dvoboja,
- datum pričetka dvoboja,
- kraj, v katerem je bil dvoboj odigran,
- datum, ko je bila odigrana posamezna partija,
- zaporedna številka partije v dvoboju,
- priimek in ime igralca z belimi figurami,
- priimek in ime igralca s črnimi figurami,
- priimek ocenjevanega igralca,
- rezultat partije.

Iz zajetih podatkov sem izgradil relacijsko podatkovno bazo, kar bom podrobneje opisal v nadaljevanju.

2.8 Podatkovna baza

Zaradi enostavnejše, učinkovitejše in preglednejše obdelave pridobljenih podatkov sem izgradil relacijsko podatkovno bazo. Entitetni model relacij med tabelami te podatkovne baze je prikazan na sliki 2.1.



Slika 2.1: Entitetni model relacijske podatkovne baze (brez atributov)

Vsak igralec je bil udeležen vsaj v enem dvoboju, le-ti so sestavljeni iz več partij, vsaka partija vsebuje eno ali več potez, posamezne poteze pa ocenjujemo pri različnih globinah preiskovanja preiskovalnega drevesa.

Tak relacijski model nam omogoča, da izvajamo analize na različnih nivojih: igralcev, dvobojev, partij in potez. Na nivoju vsakega izmed igralcev lahko npr. ugotovimo povprečno odstopanje ocen njihovih potez od potez, ki jih kot najboljše predlaga računalnik, z analizo po dvobojih lahko spoznamo, kdaj je bil kateri svetovni prvak v vzponu in kdaj v zatonu po kvaliteti njegove igre, podobne analize lahko opravljamo za posamezne dvoboje itd.

2.9 Kriteriji ocenjevanja

Že takoj na začetku je bilo jasno, da postaviti kriterije ocenjevanja, po katerih bi lahko suvereno zatrdili, kdo je bil najboljši svetovni šahovski prvak vseh časov, ne bo lahka naloga.

Pojavilo se je vprašanje, kdo je boljši igralec: tisti, ki vleče *objektivno čim boljše poteze* in v svojih partijah dela čim manjše napake in čim manj napak, ali tisti, ki s svojimi potezami, ki morda niso objektivno najboljše, na šahovnici ustvari *kompleksnejše pozicije*, v kateri se pri soočanju z v omejenem času težko rešljivih problemov bolje znajde od nasprotnika. Navsezadnje je za zmago v šahovski partiji potrebna tudi (vsaj ena) napaka nasprotnika, brez te bi se partija končala z remijem.

Pri svoji nalogi sem se skušal držati znanstvenih načel in sem zato predpostavil, da je boljši igralec tisti, ki vleče objektivno gledano boljše poteze in v svoji igri dela čim manjše napake. Vendar pa sem hkrati želel postaviti tudi kriterije, po katerih bi lahko ugotovil še nekatere druge lastnosti ocenjevanih igralcev oz. njihove igre, kot so: kompleksnost pozicij, s katerimi so se soočali oz. težavnost odločanja v njih, odstotek povlečenih najboljših potez, težnje po menjavah figur oz. po poenostavitvah pozicije itd.

Izbrani osnovni kriteriji so bili naslednji:

- povprečno odstopanje med odigranimi in predlaganimi najboljšimi potezami,
- število grobih napak,
- težavnost odločanja,
- odstotek odigranih najboljših potez,
- material tekom partije.

Pri nadaljnji obdelavi rezultatov, pridobljenih s pomočjo osnovnih kriterijev, sem poizkusil oblikovati še dodatne kriterije, s pomočjo katerih sem želel igralce oz. rezultate njihove igre še dodatno spraviti na »skupni imenovalc«. Rezultat sta naslednja *kombinirana kriterija*:

- vrednost napak v povprečno kompleksnih pozicijah,
- odstotek odigranih najboljših potez pri povprečnih odstopanjih med najboljšo in drugo najboljšo potezo.

2.9.1 Osnovni kriteriji

2.9.1.1 Povprečno odstopanje med odigranimi in predlaganimi najboljšimi potezami

Povprečno odstopanje med numerično vrednostjo potez, ki so jih svetovni prvaki odigrali v obravnavanih partijah, in med numerično vrednostjo potez, ki jih je v istih pozicijah predlagal računalnik, je bilo izbrano kot glavni kriterij za ocenjevanje.

Izračunava se po naslednji formuli:

$$\frac{\sum \text{locena najboljše poteze} - \text{ocena odigrane potezel}}{\text{št. vseh potez}}$$

Upošteva se torej povprečna *absolutna vrednost odstopanj* med ocenami pri obravnavanih potezah. Po uveljavljenem standardu za dodeljevanje ocen šahovskim pozicijam računalnik posamezne poteze ocenjene pozitivno, če gre ocena v prid belemu, in negativno, če je ocenjeno, da ima prednost igralec s črnimi figurami. Ko ima torej ocenjevani igralec črne figure, so odstopanja negativna. Upoštevanje absolutne vrednosti reši ta problem.

V primeru, da je bila odigrana najboljša poteza, je odstopanje pač enako 0, poteza pa se seveda upošteva.

Kot je bilo že pojasnjeno, sem partije ocenjeval od 12. poteze dalje. Pri upoštevanju ocen pri tem kriteriju pa sem ocenjevanje še dodatno omejil. Tako pri izračunu nisem upošteval potez, pri katerih imata tako najbolje ocenjena poteza kot odigrana poteza oceno izven območja od -2.00 do +2.00. To sem storil iz naslednjih razlogov.

- 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.

Lahko se npr. zgodi, da je najboljša poteza povezana z *žrtvijo* katere od figur, ki bi vodila do *matnega napada* na nasprotnikovega kralja, torej do hitre zmage, vendar bi bilo pri tem napadu potrebno izračunati kompleksne variante, pri katerih bi bila

možnost napake v izračunu razmeroma velika. Zato se igralec npr. raje odloči za »mirnejšo potezo«, ki prav tako vodi do zmage, pa čeprav v več potezah.

- Ko šahist ocenjuje, da je njegova pozicija *izgubljena* oz. da se niti z morebitnimi najboljšimi potezami ne more rešiti poraza, če nasprotnik ne naredi kakšne večje napake, včasih namenoma povleče potezo, za katero sicer ve, da je objektivno slabša, vendar mu daje večje *praktične možnosti*, da se reši iz težkega položaja. Lahko npr. odigra potezo, ki nastavlja skrito zanko, ali pa potezo, ki ustvari komplikacije, ki se sicer lahko ob najboljših potezah nasprotnika razpletejo v hiter poraz zanj, vendar tako poveča možnosti, da bo nasprotnik storil napako, ki lahko spremeni potek partije.

Oba navedena primera sta s stališča praktičnosti šahistov v šahovskih dvobojih povsem upravičena. Neprimerno bi bilo obsojati igralce, ko so na ta povsem legitimen način poskušali priti oz. tudi prišli do zelenega rezultata. Zato je bilo potrebno določiti mejo, ko se neka pozicija na splošno že lahko opredeli kot *dobljena* ali *izgubljena*. Crafty jemlje kot *najmanjšo odločilno prednost (minimal decisive advantage)* oceno ± 1.50 . Iz previdnosti, da se ne bi z določitvijo prenizko postavljene meje neupravičeno izgubilo preveč informacij o napakah ocenjevanih igralcev, sem po posvetovanju z nekdanjim mladinskim podprvakom Evrope, velemejstrom in reprezentantom ZRJ Draganom Šolakom, ki med drugim tudi zaključuje študij matematike v Novem Sadu, in po upoštevanju svojih lastnih izkušenj glede numeričnih ocen pri različnih šahovskih programih to mejo povečal na ± 2.00 .

Pri ocenjevanju se torej upošteva vsaka poteza, pri kateri se odigrana poteza nahaja znotraj določenega območja, omejenega z ocenama ± 2.00 . V primeru, da se najboljša poteza nahaja zunaj tega območja, odigrana pa ne, se razlika med njima upošteva kot napaka igralca, saj z odigrano potezo po naših kriterijih ni dosegel/ohranil odločilne prednosti ali pa se rešil iz pozicije, v kateri je imel odločilno prednost nasprotnik.

Namen ugotavljanja povprečnega odstopanja med odigranimi in predlaganimi najboljšimi potezami je podati čimbolj objektivno oceno o dejanski kvaliteti igre obravnavanih igralcev. Pri tem v veliki meri izhajamo iz predpostavke, da je računalnik pri svojih izračunih dejansko sposoben pravilno ugotoviti najboljše poteze, hkrati pa tudi, da so njegove numerične ocene vsaj relativno točne in da vsaj kolikor toliko pravilno odražajo razmerja med dejanskimi vrednostmi ocenjevanih potez. Ker vemo, da to ni povsem res in da Crafty ni niti najboljši obstoječi računalniški šahovski program, po vsej verjetnosti pa tudi ni boljši od vsaj nekaterih

obravnavanih igralcev, kaj šele da bi bil popoln, se upravičeno lahko vprašamo, če je sploh dovolj kompetenten za ocenjevanje svetovnih šahovskih prvakov. Odgovor na to vprašanje podajam v poglavju 2.10, na tem mestu naj omenim le, da so primerjave rezultatov pokazale, da povprečna odstopanja med odigranimi potezami in predlaganimi najboljšimi potezami računalnika, ki sem jih ugotavljal na nivoju posameznih partij, dejansko močno sovpadajo z rezultati v teh partijah, kar neposredno dokazuje tako primernost Craftyja za ocenjevalca, kakor tudi utemeljenost te metode ocenjevanja.

2.9.1.2 Število grobih napak

Pri ugotavljanju grobih napak se lahko na računalnik zanesemo z veliko mero gotovosti, kajti čeprav je pri svojih ocenah lahko nenatančen, se le redko zgodi, da bi zaradi netočnosti ocen ne bilo moč zabeležiti večjih napak oz. *spregledov* ali da bi kakšno potezo po krivem proglasil za veliko napako.

Napake sem ugotavljal podobno kot pri prejšnjem kriteriju: gre za odstopanje med ocenami odigranih potez in ocenami, ki jih kot najboljše predlaga računalnik. S to razliko, da kot grobe napake opredeljujem le poteze, pri katerih je presežen prag napake 1.00. Napaka, izražena z numerično vrednostjo 1.00, je ekvivalentna spregledu enega kmeta brez *kompensacije* oz. nadomestila zanj, kar je na ravni dvobojev za svetovnega prvaka res velika napaka. Kakor pri prejšnjem kriteriju tudi tu nisem upošteval potez, pri katerih imata tako najbolje ocenjena poteza kot odigrana poteza oceno izven območja od -2.00 do +2.00, iz že navedenih razlogov.

2.9.1.3 Težavnost odločanja (kompleksnost pozicije)

Pri določitvi kriterija za ugotavljanje težavnosti odločanja oz. *kompleksnosti pozicij* je bilo potrebno najprej razmisliti, kaj je tisto, kar vpliva na težavnost odločitev šahista.

Sprejemanje odločitev ni težko v pozicijah, kjer je na voljo več približno enako dobrih potez, saj pri odločanju ni potrebno pregledati vseh detajlov, s pomočjo katerih bi lahko določili točne ocene vsake izmed njih (podobno kot to počnejo računalniški šahovski programi), ampak lahko igralec enostavno izbere eno potezo, za katero ugotovi, da je ena izmed

najboljših in tudi v primeru majhne napake le-ta najverjetneje ne bo vplivala na končni rezultat. Še lažje se je odločiti v pozicijah, pri katerih je ena poteza očitno boljša od ostalih. Računalniški šahovski program in človek šahist pri iskanju najboljših potez ne postopata na tako povsem drugačen način, kot se to morda dozdeva. Oba iščeta v globino, recimo temu navideznega drevesa, sestavljenega iz vseh možnih potez v dani poziciji, odgovorov nanje, odgovorov na te odgovore itd. Oba se pri tem iskanju trudita, da pri tem ne bi gledala potez, ki za oceno trenutne pozicije niso relevantne, le vsak na svoj način: računalnik se pri tem opira na algoritme za učinkovito *rezanje poddreves* oz. neobravnavanje nerelevantnih delov drevesa potez, medtem ko se človek pri tem opira predvsem na svoje znanje in izkušnje. Oba sta pri iskanju časovno omejena in ne moreta iskati poljubno globoko, zato poskušata v določenem trenutku oceniti pozicijo, ne da bi preiskovala dalje. Pri tem si pomagata z delnimi ocenami na različnih globinah iskanja: računalnik pri tem pozicije ocenjuje numerično, človek pa navadno uporablja ocene »majhna prednost«, »odločilna prednost«, »nejasna pozicija« itd. Oba ne podata ocene neke pozicije, ne da bi poprej izračunala vseh *forsiranih variant* oz. serije potez z »vsiljenimi« odgovori, ki nastopijo pri menjavah figur, dajanju šahov in odgovorov nanje itd. (računalnik za ta namen uporablja *iskanje mirovanja*, kot je že bilo pojasnjeno), saj tovrstne poteze lahko neposredno vplivajo na oceno pozicije.

Med računalnikovim in človekovim postopkom iskanja najboljših potez je torej moč najti kar precej vzporednic. Zakaj je to pomembno? Pomembno je predvsem zato, ker je bilo potrebno ugotoviti, če nam lahko računalnik s svojim postopkom ocenjevanja in iskanja najboljših potez lahko poda odgovor tudi na vprašanje, kako težavno je bilo odločanje pri izbiri najboljših potez v isti poziciji za človeškega šahista.

Ideja je naslednja: neka pozicija je za opredelitev točnejše ocene in za iskanje najboljših potez težavna takrat, ko pri različnih globinah preiskovanja odkrivamo nove poteze in nadaljevanja, ki *spreminjajo oceno izhodiščne pozicije*. Igralec mora tako pregledati več nadaljevanj in pri tem iskati dovolj daleč od izhodiščne pozicije, da bi našel poteze, ki lahko odločilno vplivajo na oceno izhodiščne pozicije.

Pri ugotavljanju težavnosti odločanja oz. kompleksnosti pregledovanih pozicij sem torej primerjal ocene pri različnih globinah iskanja. Za vsako potezo sem zabeležil, kolikokrat si je računalnik na različnih globinah iskanja »premisлил«, katera je najboljša poteza. Hkrati sem pri

vsaki potezi zabeležil še vsoto razlik med ocenami najboljših in drugih najboljših potez pri globinah iskanja, ko si je računalnik premislil, katera je najboljša poteza.

Celoten postopek za zajem teh podatkov za vsako posamezno potezo bi lahko opisal takole:

```

števecSprememb := 0;           # števec postavimo na nič
vsotaOcenSprememb := 0;       # začetna vsota naj bo nič

za vsako globino, v naraščajočem vrstnem redu {
    če je globina enaka 2 {     # pri začetni globini
        prejšnjaNajboljšaPoteza := najboljšaPoteza # le zabeleži najboljšo potezo
    }
    v nasprotnem primeru {
        če prejšnjaNajboljšaPoteza ni enaka najboljšaPoteza { # našli smo novo najboljšo potezo
            števecSprememb++ # povečaj števec sprememb
                                # prištej razliko vsoti ocen sprememb
            vsotaOcenSprememb += |ocenaNajboljšePoteze – ocenaDrugeNajboljšePoteze|
            prejšnjaNajboljšaPoteza := najboljšaPoteza # zabeleži novo najboljšo potezo
        }
    }
}

```

Slika 2.2: Postopek za ugotavljanje težavnosti odločanja pri posameznih potezah

Upošteval sem torej le ocene, ko si je računalnik premislil glede najboljše poteze pri povečani globini iskanja. Slabša možnost bi bila upoštevanje razlik med ocenami najboljših potez pri različnih globinah, četudi se opredelitev najboljše poteze ne bi spremenila. Če ima namreč igralec ves čas na voljo le eno potezo, potem točna ocena te poteze niti ni pomembna za ocenjevanje težavnosti odločanja, saj igralec v pomanjkanju alternativ tako nima težke izbire.

Zakaj upoštevam razliko med najbolje ocenjeno in drugo najbolje ocenjeno potezo in ne razlike med trenutno najbolje ocenjeno in pri prejšnji globini iskanja najbolje ocenjeno potezo? Razlog je, da nova spoznanja pri povečani globini iskanja nemalokrat vodijo do večjih sprememb v ocenah potez, hkrati pa se pri povečani globini iskanja lahko pojavi več alternativnih potez, ki se po vrednosti svojih ocen lahko močno razlikujejo med seboj ali pa ne. V primeru, da so nove alternative med seboj precej izenačene (vendar se po oceni lahko precej razlikujejo od najboljše poteze pri prejšnji globini iskanja), je igralec sicer imel nekoliko težjo odločitev, ker je moral iskati do globine, pri kateri jih je odkril, vendar pri izbiri med njimi ni imel težke odločitve, saj je več kot ena možnost vodila do približno enako

dobre izbire. Če pa bi upoštevali razliko med najbolje ocenjeno novo potezo in prejšnjo najbolje ocenjeno potezo, pa bi prispevek k povečanju ocene težavnosti odločanja lahko bil kljub temu neupravičeno visok.

Globina	Najboljša poteza	Ocena	2. najboljša poteza	Ocena	Odigrana poteza	Ocena
2	Dc2	-0.09	Dc1	-0.17	Tc4	-0.75
3	Dc2	+0.24	Dc1	+0.16	Tc4	-0.45
4	Dc2	+0.08	Dc1	+0.00	Tc4	-0.73
5	Dc2	+0.35	Dc1	+0.30	Tc4	-0.44
6	Dc2	+0.07	Dc1	+0.02	Tc4	-0.72
7	Dc2	+0.57	Dc1	+0.55	Tc4	-0.43
8	Dc2	+0.72	Dc1	+0.60	Tc4	-0.42
9	Dc2	+0.96	Dc1	+0.87	Tc4	+0.12
10	Dc1	+1.30	Dc2	+1.16	Tc4	+0.10
11	Dc1	+1.52	Dc2	+1.26	Tc4	+0.33
12	Dd4	+4.46	Dc1	+1.60	Tc4	+0.28

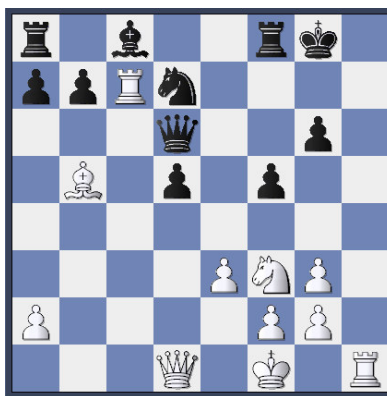
Tabela 2.4: Vrednosti atributov tabele Globina v podatkovni bazi. Nanašajo se na 20. potezo partije med Euwejem in Aljehinom, ki je podrobneje predstavljena v nadaljevanju.

Primer računanja ocene kompleksnosti pozicije bom ilustriral s pomočjo zgornje tabele. Na levi strani je navedena globina iskanja, pri kateri je računalnik prišel do izračunanih ocen, nato pa si sledijo najboljša poteza, druga najboljša poteza in poteza, ki je bila odigrana v partiji, ter njihove pripadajoče ocene.

Računalnik si je glede najboljše poteze premislil pri globini 10, vendar takrat ta sprememba ni bila tako bistvena, saj bi pri igralcu, ki bi odigral drugo najboljšo potezo, zabeležili le majhno napako z vrednostjo 0.14, kolikor znaša razlika med najboljšima potezama. Do pomembnejše ugotovitve je prišel na globini 12, ko je odkril, da bi s pomočjo poteze Dd4 igralec lahko dosegel odločilno prednost. Prispevek k skupni oceni kompleksnosti pa je bil tu precej večji: $4.46 - 1.60 = 2.86$. Le-ta tako za to potezo znaša: $0.14 + 2.86 = 3.00$, kar je veliko (povprečna ocena kompleksnosti pri potezah vseh igralcev znaša 0.335). Visoka ocena je upravičena, saj bi tudi igralec moral čez podoben miselni postopek in računati dovolj daleč, da bi pravilno izbral najboljšo potezo, hkrati pa ima razlika v odločitvah veliko težo.

To mi je tudi dalo slutiti, kako bi algoritem za izboljšanje ugotavljanje lahko še izboljšali. Spremembe izbire najboljše poteze pri različnih globinah bi lahko različno *utežili*. Tako bi imela največjo težo premišljanja na največjih globinah, kar je v skladu s tem, da mora igralec za pravilno odločitev iskati več potez naprej, kar nedvomno otežuje njegovo nalogo. Vendar

to ni povsem enostavno, saj se pojavi vprašanje, kako primerno utežiti rezultate na različnih globinah. Eden izmed vzrokov je heuristika *iskanje mirovanja*, ki jo računalnik pri našem načinu dela uporabi pri vsaki globini iskanja. Njena izklopitev pri nižjih globinah se namreč ni obnesla, saj je računalnik delal preveč napak, ki so izhajali iz problemov, ki jih prinaša *učinek obzorja*. Iskanje mirovanja namreč pogosto izenači rezultate pri različnih globinah, saj pri vseh gleda tako daleč, dokler se pozicija ne stabilizira. To pa lahko privede do tega, da rezultati utežitev ne bi odražali dejanske težavnosti odločanja igralcev.



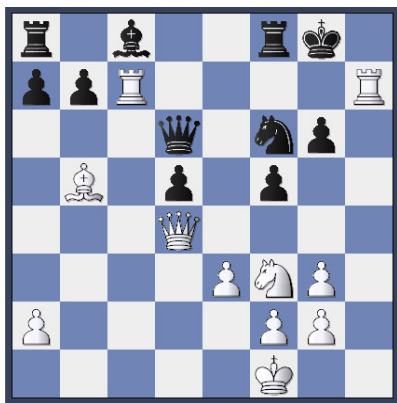
Slika 2.3: Euwe-Aljehin, dvoboj za svetovnega prvaka 1935, 14.partija. Beli, ki je na potezi, bi lahko s potezo 20.Dd4!! hitro odločil partijo v svojo korist.

Diagram na zgornji sliki se nanaša na prejšnjo tabelo. Računalnik je šele pri največji globini odkril zmagovito nadaljevanje:

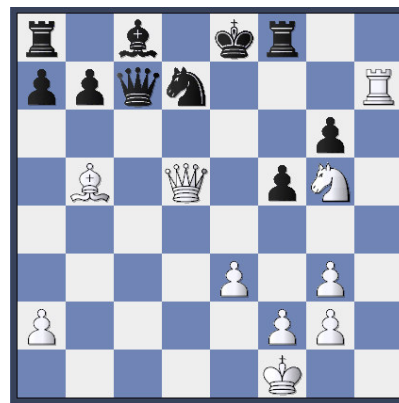
20. Dd4!! Dxc7 (ne pomaga 20... Tf6 zaradi 21.Dh4! Kf7 22.Dh8 z odločilnim napadom, niti 20... Sf6 zaradi 21.Thh7! in črni nima zadovoljive obrambe pred matom (slika 2.4, diagram levo))

21.Th8+ Kf7 22.Sg5+ Ke7 (22... Ke8 vodi po 23.Txf8+! Kxf8 24.Se6+ do izgube kraljice) **23.Th7+ Ke8** (na 23... Kd6 črni izgubi kraljico po 24.Df4+) **24. Dxd5** in šahi belemu kralju lahko mat črnemu le odložijo, ne morejo pa ga preprečiti. Končno pozicijo, ki bi tako nastala, prikazuje diagram na sliki 2.5.

Euwe je namesto poteze 20.Dd4 odigral potezo 20.Tc4, s katero je sicer ohranil majhno prednost in v nadaljevanju borbe v 41. potezi tudi dosegel zmago. Vendar smo mu pri tej potezi kljub temu zabeležili grobo napako, saj bi lahko odigral veliko boljše potezo (razlika v ocenah odigrane in najboljše poteze je kar 4.18). Za razliko od odigrane poteze bi mu prinesla odločilno prednost.



Slika 2.4: Na 20... Sf6 bi sledilo 21.Thh7!



Slika 2.5: Končna pozicija

Ugotavljal sem tudi korelacijo med težavnostjo odločanja pri posameznih potezah in napakami igralcev pri teh istih potezah. Raziskava je pokazala, da igralci v kompleksnejših pozicijah po izbranem kriteriju dejansko storijo več napak. To nedvoumno kaže na korektnost uporabljene metode in hkrati potrjuje Craftyja kot verodostojnega ocenjevalca. Rezultati te raziskave so predstavljeni v poglavju 2.10.

2.9.1.4 Odstotek odigranih najboljših potez

Kriterij *odstotek odigranih najboljših potez* sam po sebi ne pove toliko o kvaliteti igralca, kot bi to morda bilo pričakovati. V nekaterih tipih pozicij je namreč občutno lažje najti potezo, ki odstopa od ostalih v pozitivnem smislu in je torej boljša od ostalih.

Zato sem ta kriterij uporabil v kombinaciji z ugotavljanjem *odstopanja najboljše poteze od druge najboljše poteze*. Lahko namreč predpostavimo, da v pozicijah, kjer druga najboljša poteza očitno zaostaja po oceni za najboljšo potezo, le-te navadno ni tako težko izbrati kakor v nasprotnem primeru.

Za odstotek odigranih najboljših potez je bilo tako pričakovati, da bo sovpadal z odstopanjem najboljše poteze od druge najboljše poteze v istih pozicijah. Iz kakršnihkoli odstopanj od tega sovpadanja pa je že bilo mogoče sklepati o kvaliteti igralčeve igre.

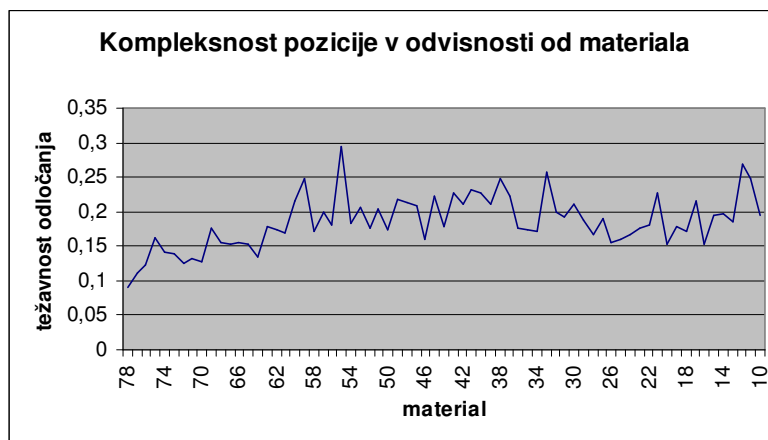
Kakor pri prvih dveh kriterijih tudi tu iz že utemeljenih razlogov nisem upošteval potez, pri katerih imata najboljše ocenjena poteza in odigrana poteza obe oceno izven območja od -2.00 do +2.00.

2.9.1.5 Material tekom partije

Namen ugotavljanja *povprečne količine materiala* oz. številsko izražene osnovne vrednosti figur na šahovnici tekom partije ni bil preverjati kvalitete igre obravnavanih igralcev, pač pa dobiti dodatne informacije o načinu njihove igre.

Predvsem je bil namen ugotoviti *težnjo posameznih igralcev po menjavah* oz. poenostavitvah pozicije. Tako sem preveril povprečno količino materiala pri posameznih potezih v partiji (od prve poteze dalje) pri vsakem igralcu posebej in primerjal odstopanja vsakega od njih s skupnim povprečjem vseh igralcev.

Želel sem preveriti hipotezo, da kompleksnost pozicij oz. težavnost odločanja *ni* povezana s količino materiala na šahovnici. Nekdo bi namreč lahko sklepal, da je težje sprejemati odločitve v pozicijah, kjer je na šahovnici več materiala in tako načeloma tudi več možnih potez (čeprav šahisti iz lastnih izkušenj vemo, da to pogosto ni res). V ta namen sem izračunal povprečno količino materiala na šahovnici pri posameznih stopnjah kompleksnosti pozicij (po intervalih) najprej pri vseh igralcih skupaj in nato še pri vsakem od igralcev posebej.



Slika 2.6: Kompleksnost pozicije v odvisnosti od materiala

Pri ugotavljanju korelacije se je jasno pokazalo, da kompleksnost pozicije ni odvisna od količine materiala na šahovnici, kar nazorno prikazuje slika 2.6.

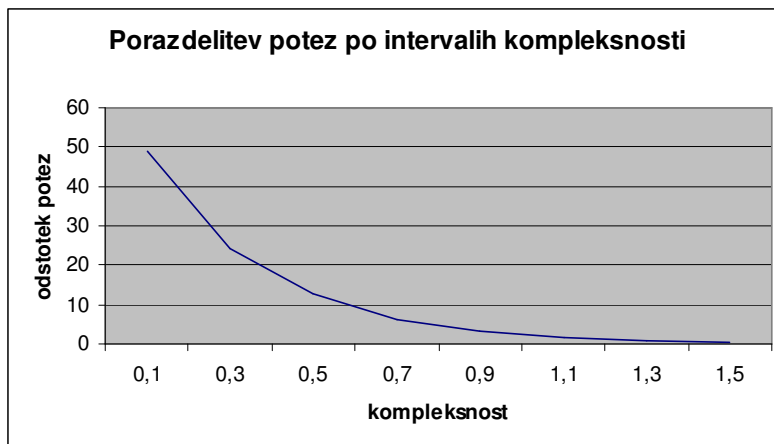
2.9.2 Kombinirani kriteriji

Osnovni kriteriji, ki sem jih izbral za ocenjevanje svetovnih prvakov, vsekakor podajajo objektivnejša merila za ugotavljanje kvalitete igralcev od zgolj statistične obdelave rezultatov njihovih partij. Vendar v svoji zasnovi ne upoštevajo, da so obravnavani igralci imeli različne *stile* igre in so torej imeli opravka z različnimi *tipi pozicij*. Tako so *pozicijski igralci* s svojim načinom igre navadno prihajali v mirnejše pozicije, v katerem so v povprečju delali manj napak kakor *taktični igralci*, ki so imeli opravka s kompleksnejšimi pozicijami, vendar so se v njih bolje odrezali od njihovih nasprotnikov in na ta način dosegali vrhunske rezultate.

Pojavlja se vprašanje, kako bi se pozicijski igralci odrezali v pozicijah taktičnih igralcev in obratno. Na to vprašanje v primeru obravnavanih svetovnih prvakov seveda nikoli ne bo mogoče natančno odgovoriti. Nikoli ne bomo izvedeli npr. kako uspešen bi bil Capablanca, ki velja za tipičnega pozicijskega igralca, v pozicijah Talja, ki se je ravno s pomočjo uspešnega reševanja zapletov v kompleksnih taktičnih pozicijah prebil v sam svetovni vrh. Vendar pa lahko s pomočjo simulacij, ki so bile uporabljene pri kombiniranih kriterijih, vsaj približno podamo odgovore na ta in podobna vprašanja.

Pri svojem delu sem se tako znašel pred novim izzivom: kako definirati *stil igre*? Kaj je tisto, kar loči pozicijske igralce od taktičnih igralcev? Razmislek je vodil do naslednjega sklepa: osnovna delitev na ta dva tipa igralcev je opredeljena predvsem s kompleksnostjo pozicij, kot smo jo bili definirali. Taktični igralci so namreč navadno morali pregledati precej več nadaljevanj, da bi se dokopali do najboljših potez in ocenili njihove vrednosti, in so njihove odločitve navadno imele večjo težo kot pri pozicijskih igralcih. Seveda pa obstajajo tudi univerzalni igralci, ki jih ne bi mogli opredeliti samo v eno od obeh skupin.

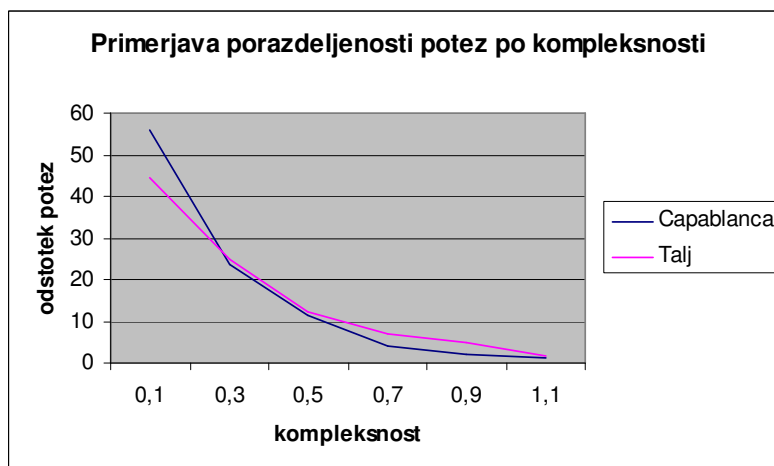
V namen ugotavljanja stilov igralcev sem razdelil izračunane stopnje kompleksnosti na več intervalov in pri vsakem igralcu preveril, kolikšen odstotek odigranih potez pripada kateremu intervalu. Naslednja slika predstavlja povprečno porazdelitev potez po teh intervalih pri vseh igralcih skupaj.



Slika 2.7: Povprečna porazdelitev potez po intervalih kompleksnosti pri vseh igralcih

Igralci so torej v povprečju večinoma imeli opravka z manj kompleksnimi pozicijami.

Slika 2.8 prikazuje tovrstni krivulji igralcev, ki sta v tem pogledu najbolj odstopala. To sta ravno zgoraj omenjena igralca Talj in Capablanca.



Slika 2.8: Primerjava porazdeljenosti potez po kompleksnosti med Taljem in Capablanc

Rezultat torej nazorno potrjuje našo hipotezo, da je pozicijski igralec Capablanca večkrat imel opravka z manj kompleksnimi pozicijami kot taktični igralec Talj.

Omenjene ugotovitve so mi omogočile, da sem lahko primerjal igralce med seboj glede na različne stile igre. Glavno merilo pa je predstavljalo povprečje stilov vseh igralcev.

2.9.2.1 Vrednost napak v povprečno kompleksnih pozicijah

Namen tega kriterija je bil *simulirati povprečne vrednosti odstopanj* med odigranimi potezami in najboljšimi potezami pri uravnoteženi kompleksnosti pozicij. Z drugimi besedami, želel sem ugotoviti vrednost napak, ki bi jih igralci delali v pozicijah, kjer bi bil delež potez v vseh intervalih kompleksnosti pri vseh igralcih enak in bi tako predstavljale povprečno kompleksne pozicije glede na partije njih vseh.

Uporabljen postopek je bil naslednji:

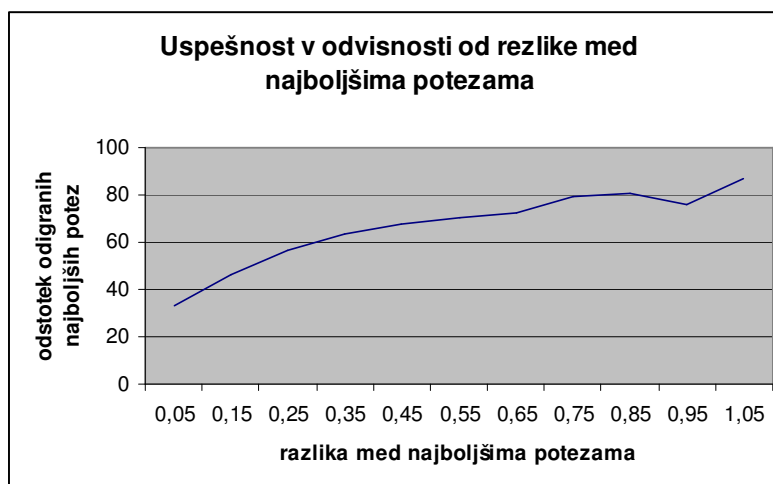
- pri vsakem igralcu izračunamo, kakšno *povprečno napako* je naredil v določenih intervalih kompleksnosti,
- zabeležimo porazdelitev potez po intervalih kompleksnosti za vsakega igralca posebej,
- izračunamo *povprečno kompleksno partijo*: izmerimo povprečno zastopanost potez v vsakem intervalu kompleksnosti, pri čemer upoštevamo vse igralce enakovredno, ne upoštevamo torej števila potez v partijah, ampak deleže zastopanosti le-teh v vsakem posameznem intervalu,
- ugotovimo, kakšno napako pri posamezni potezi bi v povprečju naredil igralec, če bi igral v stilu, ki ga predstavlja tako definirana povprečno kompleksna partija.

Na podoben način lahko ugotovimo *povprečne napake pri različnih stilih igranja*. Kot smo določili povprečno kompleksno partijo za vse igralce, jo lahko tudi za vsakega posameznega igralca posebej. Ko nato ugotavljamo, kakšno napako pri posamezni potezi bi v povprečju naredil igralec pri spremenjenem stilu igranja, za uravnoteženje namesto povprečno kompleksne partije vseh igralcev uporabimo povprečno partijo določenega igralca, npr. Talja ali Capablance, ki sta tipična predstavnika različnih stilov.

Ker imajo različni igralci različno zastopanost potez pri posameznih intervalih kompleksnosti, se pri tem ustrezno spremenijo tudi rezultati vsakega igralca. Pomembna pri tem je uspešnost igralca v intervalih kompleksnosti, ki dobijo po spremenjeni utežitvi večjo težo.

2.9.2.2 Odstotek odigranih najboljših potez pri povprečnih odstopanjih med najboljšo in drugo najboljšo potezo

Rezultati so pokazali, da je odstotek odigranih najboljših potez v tesni korelaciji z razliko v ocenama med najboljšima potezama v poziciji. Večja ko je bila ta razlika, uspešnejši so bili igralci pri izbiri najboljše poteze. Naslednja slika nazorno prikazuje omenjeno korelacijo.



Slika 2.9: Odstotek odigranih najboljših potez v odvisnosti od razlike med najboljšima potezama. Izračunana Spearmanova korelacija znaša $\rho=0.97$ ($p=0,0000051$).

Korelacija je logična, saj večje kot je odstopanje med najboljšima potezama, večjo napako bi predstavljalo ne odločiti se za najboljšo potezo, in pri svetovnih prvakah je seveda razumljivo, da je naraščanje krivulje na grafu še toliko bolj izrazito. Pri slabših igralcih bi upravičeno lahko pričakovali, da ta krivulja ne bi potekala tako visoko.

Rezultati pri ugotavljanju odvisnosti odstotka odigranih najboljših potez in razlike med najboljšima potezama nam torej lahko dajo informacije o kvaliteti posameznih igralcev. Pokaže se, da krivulje pri različnih igralcih lahko precej varirajo. Vse to pa je bila osnova za postavitev kriterija, s pomočjo katerega bi neposredno dobili informacije o kvaliteti igralcev.

Postopek pri tem kombiniranem kriteriju je bil sledeči:

- pri vsakem igralcu izračunamo delež potez, pri katerih smo ugotavljali odstotek odigranih najboljših potez, v določenih intervalih razlik med najboljšima potezama,
- zabeležimo povprečne vrednosti razlik med najboljšima potezama po posameznih intervalih pri vsakem igralcu posebej,

- izračunamo povprečne vrednosti razlik med najboljšima potezama po posameznih intervalih pri vseh igralcih, pri čemer upoštevamo vse igralce enakovredno,
- ugotovimo, kako bi se posamezni igralci izkazali pri povprečni porazdelitvi vrednosti razlik med najboljšima potezama.

Rezultat tega kriterija je pričakovani odstotek potez, ki bi jih pozamezni igralci odigrali pri tej povprečni porazdelitvi.

2.10 Verodostojnost Craftyja kot ocenjevalca

2.10.1 Crafty kot ocenjevalec napak

Crafty je vsekakor najprimernejši od vseh računalniških šahovskih programov, ki sem jih imel na voljo za uresničitev zastavljenih ciljev. Potreboval sem namreč program odprte kode, ki sem ga lahko preprogramiral, kot je bilo to potrebno za naš namen, in Crafty je po moči z naskokom najboljši med vsemi programi te vrste. Hkrati pa je tudi na splošno eden izmed najboljših računalniških šahovskih programov na svetu.

Kljub temu je vse od samega začetka bil pomemben odgovor na vprašanje, ali je Crafty dovolj primeren program za ocenjevanje svetovnih šahovskih prvakov. Današnji računalniški šahovski programi namreč še vedno ne veljajo za boljše od vrhunskih šahovskih velemejstrov in najverjetneje Crafty ni boljši vsaj od nekaterih obravnavanih svetovnih prvakov.

Vendar kljub vsemu obstaja kopica argumentov, ki jih je moč navesti v prid računalniškim šahovskim programom kot ocenjevalcem:

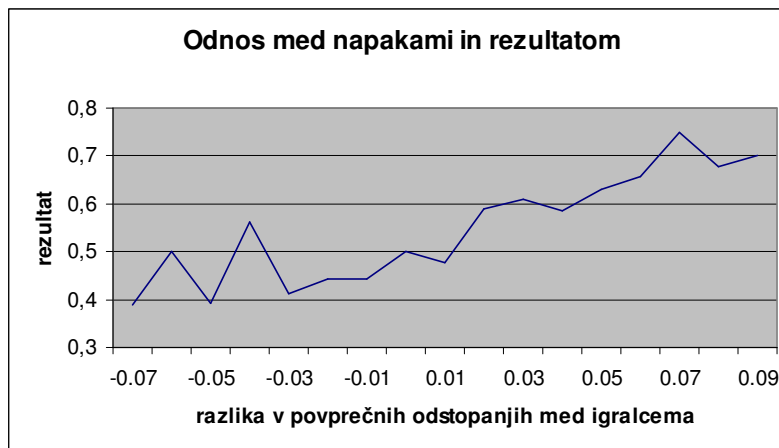
- Pri ocenjevanju pozicij navajajo natančne numerične ocene, ki so za naš namen bile neprimerno bolj uporabne kot tipične ocene ljudi šahistov, kot so navadno »majhna prednost belega«, »izenačena pozicija« itd.
- Pri ocenjevanju se ves čas držijo istih pravil in so zato pri ocenah veliko bolj konsistentni od ljudi, prav tako tudi niso pod vplivom čustev in drugih motečih dejavnikov, ki so značilni za ljudi.

- Izredno dobri so pri ocenjevanju *taktičnih pozicij*, kjer je prisotnega veliko računanja in so napake pri odločanju pri ljudeh šahistih navadno večje, za slabše pa veljajo v *pozicijski igri*, kjer so razlike med ocenami najboljših potez običajno precej manjše. Torej morebitne napake v računalnikovih ocenah v našem primeru niso imele tako velikega vpliva na končne rezultate.
- Najboljši šahisti so dandanes razmeroma uspešni proti računalniškim programom predvsem zato, ker vedo, da je njihov nasprotnik računalnik, in zato prilagodijo svojo igro tako, da kar najbolje izkoristijo omenjene slabosti računalnikov pri pozicijski igri. Pri analizi partij pa se večina šahistov najbolj zanese prav na računalniške programe.

Pri utemeljevanju Craftyja kot verodostojnega ocenjevalca sem se skušal čimbolj držati znanstvenih načel. V tem smislu je bilo bistveno predvsem ugotoviti, v kolikšni meri razlike med ocenami odigranih in najbolje ocenjenih potez dejansko sovpadajo z rezultati v partijah. V skladu z našo hipotezo, da rezultat ne odraža povsem dejanske kvalitete igre (v nasprotnem primeru bi bilo dovolj dobro merilo za ugotavljanje kvalitete igralcev enostavno primerna statistična obdelava njihovih rezultatov), je bilo pričakovati, da omenjene razlike v ocenah ne bodo povsem sovpadale z rezultati. Kljub temu pa bi se za ugotovitev verodostojnosti Craftyja moralo pokazati, da obstaja statistično značilna korelacija med tema dvema spremenljivkama.

V ta namen sem izvedel dve raziskavi. Pri prvi sem obravnaval partije, kjer so se svetovni prvaki medsebojno soočili, in primerjal sovpadanje razlik med njihovimi povprečnimi napakami z rezultati partij. Druga raziskava je bila obsežnejša, s Craftyjem sem še dodatno preanaliziral številne partije različnih igralcev z mednarodnim (FIDE) ratingom in ugotavljal korelacijo med ratingom in povprečno napako.

Prva od omenjenih dveh raziskav je torej obsegala medsebojne partije svetovnih prvakov (skupaj 523 partij). Pri vsaki od njih sem ugotovil povprečno odstopanje med ocenami odigranih potez in ocenami potez, ki jih je kot najboljše predlagal računalnik, za enega igralca in jo primerjal z istim povprečnim odstopanjem pri drugem igralcu (s stališča prvega, torej je razlika lahko bilo pozitivno ali negativno število) ter zabeležil rezultat partije. Nato sem te razlike v odstopanjih med obema igralcema razdelil po intervalih in pri vsakem od teh intervalov ugotavljal uspešnost prvega igralca.



Slika 2.10: Odnos med povprečnimi odstopanji med obema igralcema v posameznih partijah in rezultatom. Izračunana Spearmanova korelacija znaša $\rho=0.89$ ($p=0.0000020$).

Vrednost na abscisi predstavlja *povprečno razliko v napakah med dvema igralcema*. Pozitivna razlika je v korist prvega igralca (naključno izbranega od obeh), kar pomeni, da je v povprečju delal manj napak oz. manjše napake od nasprotnika, negativna razlika pa pomeni, da je prvi igralec v povprečju naredil več napak oz. večje napake kot njegov nasprotnik. Na navpični osi pa so podatki o *uspešnosti obravnavanega igralca* v teh medsebojnih partijah. Vsaka partija je bila na ta način obravnavana le enkrat, tako da v rezultatih ni podvajanj.

Iz grafa je razvidno, da so izračunane napake v igri obeh igralcev povezane z njuno rezultatsko uspešnostjo. To pa tudi dodatno potrjuje korektnost poprej izračunanega povprečnega odstopanja med ocenami odigranih potez in potez, ki jih je kot najboljše ocenil Crafty, ki sem ga uporabil kot enega izmed glavnih kriterijev za ugotavljanje kvalitete obravnavanih igralcev.

Dejstvo, da razlika med izračunanimi povprečnimi napakami dveh igralcev, ki sta se soočila v medsebojnih partijah, ne sovпада povsem z njuno rezultatsko uspešnostjo v teh partijah, potrjuje našo hipotezo, da rezultat ne odraža povsem kvalitete igre šahistov in da za njihovo medsebojno primerjavo ni dovolj zgolj statistična obdelava rezultatov. Kajti čeprav je v šahu dosti verjetno, da bo zmagal igralec, ki dela manjše napake in manj napak, temu ni vedno tako, saj lahko partijo odloči že ena sama napaka, ki v našem primeru po numerično izraženih vrednostih morda ne odtehta več manjših napak nasprotnika, vendar je za potek partije lahko usodnejša.

Namen druge raziskave je bil ugotoviti povezanost med šahovsko močjo igralcev in izmerjeno povprečno napako oz. preveriti, če boljši igralci v partijah dejansko naredijo v povprečju manj napak. Najboljše (čeprav ne popolno) merilo za ugotavljanje moči šahistov je rating Svetovne šahovske organizacije (FIDE). Obračunava se vsake tri mesece, v obračun pa so zajete praktično vse partije z mednarodnih turnirjev, državnih prvenstev itd. povsod po svetu (pri čemer se ne upoštevajo turnirji s skrajšanim časom za razmišljanje). Mnogim šahistom rating predstavlja glavni motivacijski dejavnik za udeleževanje na šahovskih turnirjih. Lestvica FIDE ratingov se je še do nedavnega pričela pri 2000 točkah, šele pred kratkim so to mejo spustili na 1800 točk. Najvišje ratingirani aktivni igralec na lestvici, objavljeni julija 2005, je Viswanathan Anand z 2788 točkami (Gari Kasparov je sicer prvi z 2812 točkami, vendar je že pred objavo lestvice naznanil prenehanje igranja na turnirjih).

S Craftyjem sem preanaliziral 1450 partij, ki jih je v obdobju med letoma 2000 in 2005 odigralo 61 aktivnih igralcev šaha, s FIDE ratingom v razponu od 2000 do 2800 ratingovskih točk. Pri izboru sem se soočil z naslednjimi problemi:

- FIDE ratingi se spreminjajo zelo počasi in pogosto ne odražajo dejanske šahovske moči igralcev oz. njihove trenutne forme,
- ratingi večine igralcev močno varirajo, igralcem je tako zelo težko določiti pravo moč,
- performansi (izračunani po formuli FIDE za njihovo računanje) igralcev na turnirjih le redko ustrezajo njihovim dejanskim ratingom, njihovi rezultati so pogosto nad ali pod glede na rating pričakovanim rezultatom.

Spričo teh razlogov in ker mi niti čas niti strojna oprema nista dopuščala analize še veliko večjega števila partij, je bilo potrebno izbrati čimbolj reprezentativno množico partij oz. partije, pri katerih je rating čimbolj ustrezal dejanski moči igralca, takrat ko je posamezna partija bila odigrana.

V ta namen sem izbral igralce, katerih rating je v ocenjevanem obdobju čim manj variral. Pri izbiri mi je pomagal šahovski mednarodni mojster in diplomirani novinar Marko Tratar, v pomoč pa nama je bila tudi uradna stran FIDE, kjer so na voljo grafi, ki prikazujejo gibanje ratingov posameznih igralcev v zadnjih letih. Pri izboru igralcev sta bila kriterija še število partij, konstantnost ratinga namreč ni smela biti posledica neigranja, in borbenost, namreč pri igralcih, ki pogosto hitro remizirajo partije, so te partije oz. vsaj njihov pretežni del pogosto

pripravljene v domačih analizah (to velja predvsem za velemojstre) in bi kaj hitro lahko izpadlo, da bi bile odigrane brez napak, kar pa ne bi bil realen odraz igralčeve moči.

Za vsakega igralca je bilo nato potrebno izbrati za njegov rating (ki naj bi pri izbranih kandidatih z avtorjeve stabilnosti zares bil dober pokazatelj njihove moči) reprezentativne partije (tipično od 20 do 25 partij). To pomeni partije, za katere je čimbolj verjetno, da jih je igralec res odigral z močjo, ki jo odraža njegov rating. Zakaj ne bi bil dober kar izbor naključnih 25 partij vsakega od igralcev? Tudi pri izbranih stabilnih igralcih je namreč bilo moč opaziti precejšnja nihanja. Kljub relativno stabilnim ratingom je le težka moč najti turnirje, ki bi po performansih igralca dobro ustrezale njegovemu ratingu. Reprezentativne partije pa naj bi bile prav takšne.

Performansi se izražajo v ratingovskih točkah. Pri njihovem računanju sem uporabljal program Chessbase, tako da so izračunani v skladu z uradno formulo FIDE. Računanje performansov ni smiselno za posamezne partije, saj bi igralec potemtakem odigral partijo s svojo močjo le, če bi remiziral proti po ratingu sebi enakemu igralcu. Veliko bolj je smiselno računanje performansov igralcev na posameznih turnirjih. Turnirji tipično vsebujejo vsaj 9 partij, za igralca pa lažje trdimo, da je bil na nekem turnirju ves čas v približno enaki formi, kot pa bi to lahko trdili za več zaporednih turnirjev, ki si lahko sledijo v razmaku tudi do par mesecev ali celo več oz. za skupino naključno izbranih partij. Za posamezne turnirje je značilno še, da ima igralec ves čas enak rating. Na dlani je torej bilo, da je smiselno izbrati partije s posameznih turnirjev, pri katerih performansi igralca, izraženi v ratingovskih točkah, čimbolj ustrezajo njegovemu ratingu (ki naj bi tudi sicer pri izbranih kandidatih bolje izražal njihovo dejansko moč, kot pri ostalih igralcih).

Posamezni igralec je lahko zastopal tudi več različnih ratingov, če je bil njegov rating na posameznih turnirjih, s katerih so bile izbrane njegove partije, različen. Pomembno pa je bilo, da je za vsako skupino partij z istim ratingom, torej za vsako skupino partij z istega turnirja (oz. z več turnirjev, kjer je ocenjevanj igralec imel isti rating), veljalo, da njihov skupni performans čimbolj ustreza igralčevemu ratingu na tem turnirju oz. da naj bi izbrane partije s posameznega turnirja vse skupaj (ne pa vsaka posebej) predstavljale približek kvalitete igre za izbrani rating. To sem kasneje tudi upošteval in sem tako izračunal povprečno napako za vsak rating posebej. Tako sem dobil pare »rating – povprečna napaka«, pri čemer so bili ratingi zastopani na celotnem intervalu od 2000 do 2800 ratingovskih točk.

Pri pregledovanju rezultatov sem ugotovil, da še posebej pri višje ratingiranih igralcih včasih prihaja do precejšnjih razlik med izmerjenimi povprečnimi napakami igralcev s podobnimi ratingi. Preveril sem, kaj bi lahko bili vzroki za ta odstopanja in v ta namen natančneje pregledal obravnavane partije pri takšnih primerih. Ugotovil sem naslednje.

- Izbrane partije so bile odigrane bodisi na odprtih bodisi na zaprtih turnirjih. Na zaprtih turnirjih igralci že vnaprej poznajo svoje nasprotnike in se nanje lahko bolje ustrezno pripravijo, tako da je kvaliteta igre pri teh tipih turnirjev navadno večja.
- Igralci se tudi razlikujejo po stilu. Npr. tisti, ki igrajo pretežno odprte turnirje, so za doseg boljšega rezultata prisiljeni tudi več tvegati, saj je za dober rezultat na odprtih turnirjih potrebno večje število točk kot pri zaprtih turnirjih, kjer sta lahko npr. že dve zmagi ob samih remijih v devetih partijah dovolj za prvo mesto na turnirju.
- Igralci so igrali proti različno močnim nasprotnikom. Tisti, ki so igrali proti slabšim nasprotnikom, so se v povprečju manjkrat znašli pred težjimi izzivi in so zato v povprečju delali manj napak. (S stališča performansov lahko npr. zmaga proti slabšemu nasprotniku pomeni enak rezultat kot neodločen rezultat proti boljšemu nasprotniku.).

Medsebojno povezanost obeh spremenljivk sem ugotavljal s Spearmanovim koeficientom korelacije. Izračunana korelacija znaša $\rho=0,42$ z značilnostjo $p=0.000053$. Korelacija med obema spremenljivkama se je torej nedvoumno pokazala, prav gotovo pa bi bila še večja, če ne bi bilo zgoraj naštetih razlogov. Naj še dodam, da pri ocenjevanju svetovnih prvakov omenjeni razlogi ne pridejo v poštev, saj so vsi igrali isti tip »turnirja« (dvoboj namreč), vsi so se dobro pripravili na svoje nasprotnike (tipično priprave na dvoboj za svetovnega prvaka trajajo tudi po več mesecev), prav tako pa je tudi vsak od njih igral proti močnim nasprotnikom.

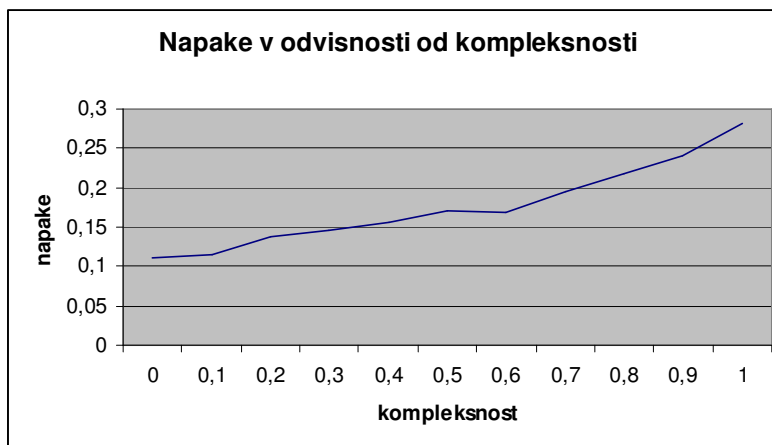
Pomemben argument v prid Craftyju kot ocenjevalcu je tudi dejstvo, da smo opravili analizo zelo velikega števila potez, kar zmanjšuje pomen morebitnih posameznih napak.

Nadaljnje raziskave, ki bi lahko z večjo mero natančnosti pokazale, s kolikšno mero lahko pri ocenjevanju kvalitete igre v šahovskih partijah zaupamo tako Craftyju kot tudi računalniškimi programom na splošno, bi lahko bile usmerjene v primerjavo odstopanj ocen pri še večjih globinah iskanja in njihov vpliv na končno oceno (I) ter primerjava ocen različnih najboljših

šahovskih programov (II). Prve raziskave sam v omejenem času in s strojno opremo, ki sem jo imel na voljo, nisem imel možnosti izvesti, enako velja tudi za drugo možnost, saj izvorna koda današnjih najboljših programov na svetu ni na voljo, to pa bi bil potreben pogoj, da bi jih lahko ustrezno spremenil oz. določil maksimalno fiksno globino iskanja, kot sem to lahko storil pri Craftyju in česar komercialne različice pri vrhunskih računalniških šahovskih programih ne ponujajo.

2.10.2 Crafty kot ocenjevalec kompleksnosti pozicij

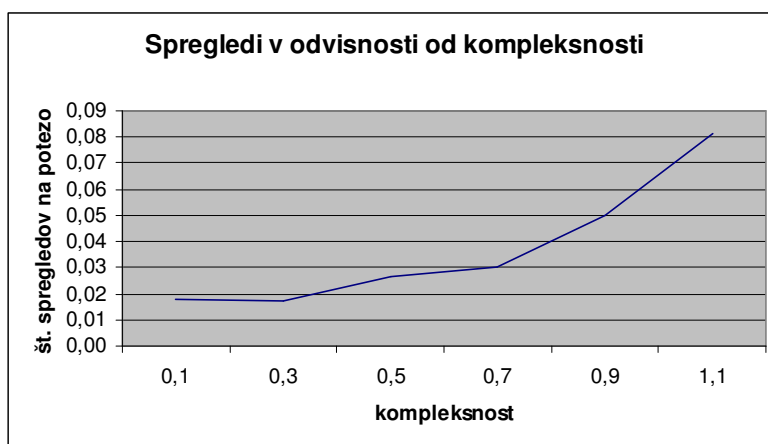
Zanimala me je tudi verodostojnost Craftyja kot orodja za ugotavljanje kompleksnosti pozicij. V ta namen sem ugotavljal korelacijo med napakami šahistov in stopnjo kompleksnosti pozicije. Zabeležil sem napake obravnavanih igralcev, izraženih z odstopanjem ocen odigranih potez od ocen najboljših potez, pri različnih stopnjah (intervalih) kompleksnosti in nato za vsak interval izračunal povprečno oceno napake. Rezultati so prikazani na sliki 2.11.



Slika 2.11: Graf napak igralcev pri različnih intervalih kompleksnosti. Izračunana Spearmanova korelacija znaša $\rho=0.99$ ($p=0,0000000038$).

Rezultat je logičen in nikakor ni nepričakovan. Igralci so najmanj napak delali v preprostih pozicijah, z naraščajočo kompleksnostjo pozicij pa vse več. Crafty se je torej tudi v tem primeru izkazal kot verodostojen ocenjevalec, metoda za ugotavljanje kompleksnosti pozicij oz. težavnosti odločanja v njih pa za korektno.

Želel sem tudi preveriti hipotezo, da so igralci v bolj kompleksnih pozicijah naredili več grobih napak oz. spregledov. Spregled je definiran kot napaka, pri kateri odstopanje od najboljše poteze presega numerično vrednost 1.00, kar je ekvivalentno izgubi enega kmeta brez nadomestila oz. *kompensacije* zanj. Postopek pri tem je bil podoben zgoraj opisanem ugotavljanju odvisnosti med napakami šahistov in stopnjo kompleksnosti pozicije. Zabeležil sem spreglede obravnavanih igralcev pri različnih stopnjah (intervalih) kompleksnosti in nato za vsak interval izračunal povprečno število spregledov.



Slika 2.12: Graf grobih napak igralcev pri različnih intervalih kompleksnosti. Izračunana Spearmanova korelacija znaša $\rho=0.98$ ($p=0,0000015$).

Rezultat je potrdil našo hipotezo, da je v bolj kompleksnih pozicijah v partijah bilo storjeno več grobih napak. Naraščajoča krivulja na zgornjem grafu to nazorno prikazuje.

2.11 Metode dela s tehničnega vidika

Pri svojem delu sem uporabljal dva programska jezika:

- C: preprogramiranje Craftyja,
- Perl: obdelava rezultatov, polnjenje relacijske podatkovne baze z obdelanimi podatki, dodatni izračuni za potrebe prezentacije rezultatov.

Pri izgradnji relacijske podatkovne baze je bil uporabljen Microsoft Access, nekatere dodatne statistične obdelave sem opravil v programskem okolju Mathematica, grafe in tabele pa sem ustvaril s pomočjo programskega produkta Microsoft Excel.

Delo je bilo razdeljeno v več faz:

1. spoznavanje z delovanjem računalniškega šahovskega programa odprte kode Crafty in preprogramiranje programa za naše potrebe,
2. poganjanje spremenjenega programa na računalnikih, ki sem jih v ta namen imel na voljo (vseh skupaj 36) in arhiviranje tekstovnih datotek z vsebovanimi analizami,
3. načrtovanje in izgradnja relacijske podatkovne baze, obdelava pridobljenih podatkov in njihovo umeščanje v podatkovno bazo,
4. delo s podatkovno bazo, programiranje skript za pridobivanje informacij iz nje ter za njihovo nadaljnjo obdelavo za potrebe prezentacije,
5. izdelava grafov in tabel ter izračuni različnih statističnih podatkov.

Pri preprogramiranju Craftyja je bil glavni izziv omejiti njegovo iskanje ne časovno, pač pa na iskanje do vnaprej določene globine iskanja oz. spremeniti program tako, da bo iterativno povečeval *fiksno globino iskanja*, dokler ne doseže *maksimalne fiksne globine*. Pri tem je bilo pomembno, da spremembe ne bi kakorkoli vplivale na njegove ocene oz. način ocenjevanja. Hkrati je bilo potrebno poskrbeti za sprotne, čimbolj informativne izpise.

Pri obdelavi rezultatov se je Perl izkazal kot odlično orodje. Zaradi razširjenega nabora *regularnih izrazov* (*regular expressions*) je ta programski jezik res izjemno primeren za delo s tekstovnimi datotekami, prav tako pa so mi prišle zelo prav njegove odlične knjižnice za delo s podatkovnimi bazami.

S tehničnega vidika nisem imel kakšnih posebnih težav. Pri delu je bila potrebna izjemna natančnost, saj bi že vsaka majhna napaka lahko imela velike posledice za končne rezultate, zato sem vse sprotne rezultate skrbno preverjal. Pri tem mi je v veliki meri prišlo prav znanje šaha in poznavanje lastnosti igre svetovnih šahovskih prvakov. Za preprečevanje napak in njihovo odpravljanje je namreč nemalokrat bilo potrebno pregledovati šahovske pozicije in preverjati njihove ocene, pa tudi zaznati morebitna odstopanja od pričakovanih rezultatov ter preveriti njihove vzroke.

3

Predstavitev rezultatov

3.1 Napake igralcev

3.1.1 Povprečna napaka pri vsaki potezi

Merjenje povprečne napake pri odigranih potezah je bilo glavno merilo za ugotavljanje kvalitete igre svetovnih prvakov. Računali smo odstopanja med odigranimi in predlaganimi najboljšimi potezami, njihovo povprečje glede na posamezno potezo pa predstavlja končni rezultat, ki je prikazan v tabeli 3.1.

Igralec	Povprečna napaka
Capablanca	0,1008
Kramnik	0,1058
Karpov	0,1275
Kasparov	0,1292
Spaski	0,1334
Petrosjan	0,1343
Lasker	0,1370
Fischer	0,1383
Aljehin	0,1402
Smislov	0,1403
Tal	0,1492
Botvinik	0,1581
Euwe	0,1777
Steinitz	0,2300

Tabela 3.1: Povprečna odstopanja med odigranimi in predlaganimi najboljšimi potezami

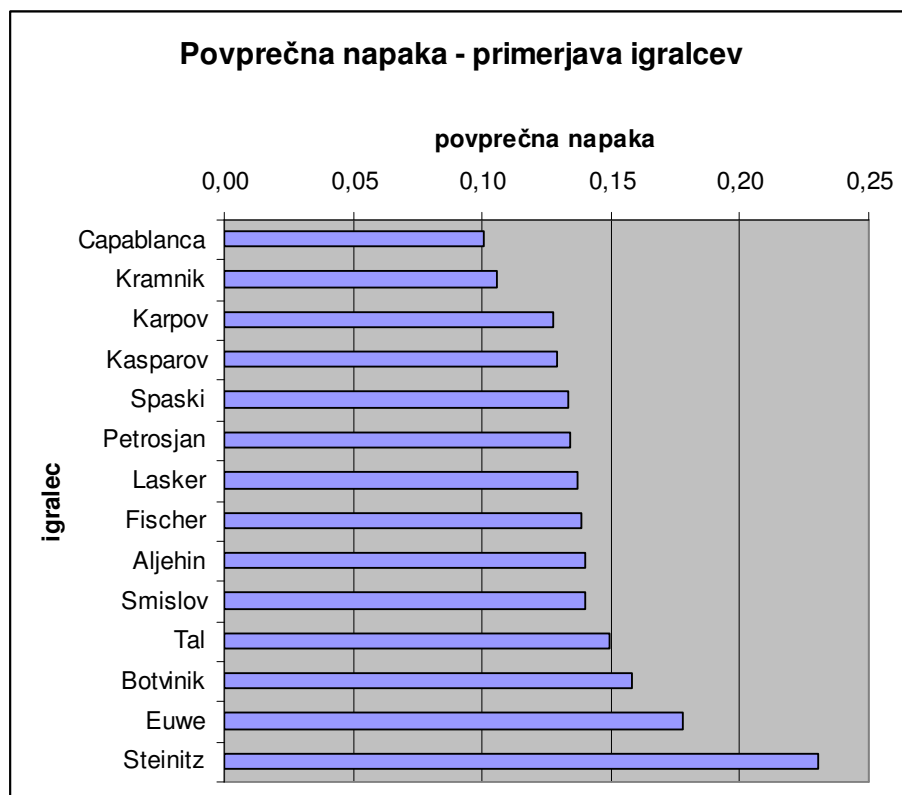
Povprečna napaka vseh igralcev znaša 0,1449. Pri njenem izračunu sem upošteval, da je pri vsakem od njih bilo ocenjevano različno število potez. Vseh ocenjevanih potez je bilo 37285.

Prvo mesto Kubanca Capablance po eni strani vsekakor *je* presenetljivo, saj gre za igralca, ki je bil na šahovskem prestolu med leti 1921 in 1927, torej daleč pred prihodom računalnikov in

ogromnih baz šahovskih partij, ki današnjim šahistom služijo za šahovsko izpopolnjevanje in pripravo na nasprotnike, prav tako pred izidom številnih odličnih knjig o šahu, ki so bile v pomoč naslednjim generacijam. Capablanca je živel v času, ko se je o šahu vedelo res malo v primerjavi z današnjim, medtem ko je teorija otvoritev, ki je zaživela kasneje s prihodom Botvinika in je dandanes že izredno razvita, takrat bila še povsem v povojih.

Po drugi strani pa je Capablanca znan kot miren pozicijski igralec, ki je še posebej slovel po enostavnosti svojih partij in po svoji izjemni natančnosti, zato je nedvomno bil kandidat za sam vrh po tem kriteriju.

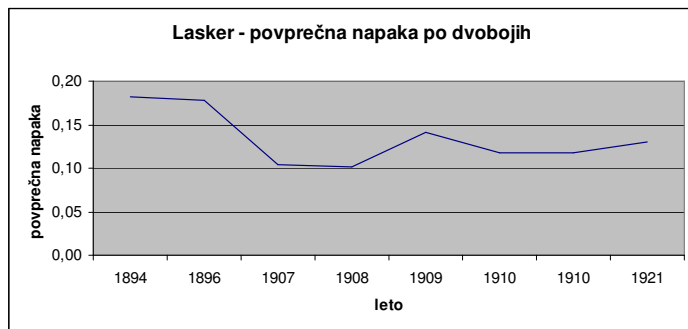
Potrdila se je hipoteza, da so manj in manjše napake delali *pozicijski igralci*, ki so dosegali uspehe predvsem zaradi natančnosti v mirnih pozicijah, medtem ko so *taktični igralci* v spodnji polovici lestvice. Povsem na dnu je Steinitz, ki je živel še v 19. stoletju in je v njegovi igri še prisoten pridih *romantičnega obdobja* v šahu, ki še posebej slovi po številnih taktičnih zapletih v partijah.



Slika 3.1: Grafični prikaz povprečnih napak igralcev

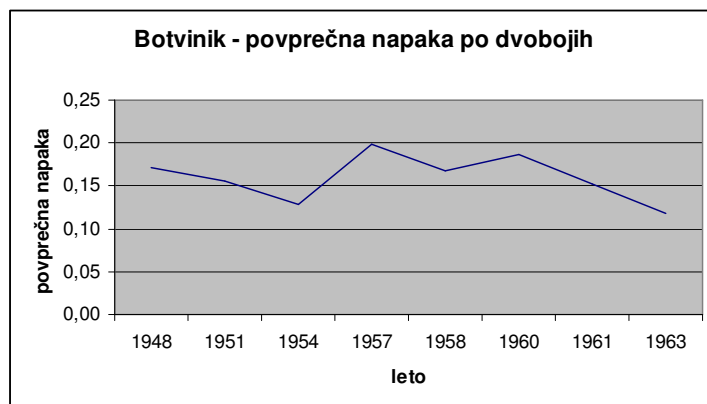
3.1.2 Primerjava igralcev po dvobojih

Pri nekaterih igralcih je bilo opaziti spreminjanje kvalitete njihove igre skozi različna obdobja. Še posebej zanimivi v tem pogledu so bili Lasker, Botvinik in Kasparov, ki so bili udeleženi v največ dvobojih za svetovnega prvaka (vsi po 8-krat), v daljših časovnih obdobjih in proti različnim nasprotnikom.



Leto	Povprečna napaka
1894	0,1817
1896	0,1784
1907	0,1044
1908	0,1007
1909	0,1414
1910	0,1173
1910	0,1184
1921	0,1299

Slika 3.2: Laskerjeva povprečna napaka na potezo v različnih dvobojih. Opazimo lahko, da je pri njegovih prvih dveh dvobojih, ki ju je odigral proti Steinitzu, povprečna napaka z naskokom največja. Hkrati pa lahko vidimo, da je v dvoboju za svetovnega prvaka, ki ga je leta 1908 odigral proti Tarraschu, njegova povprečna napaka bila manjša od skupne povprečne napake pri Capablanci!



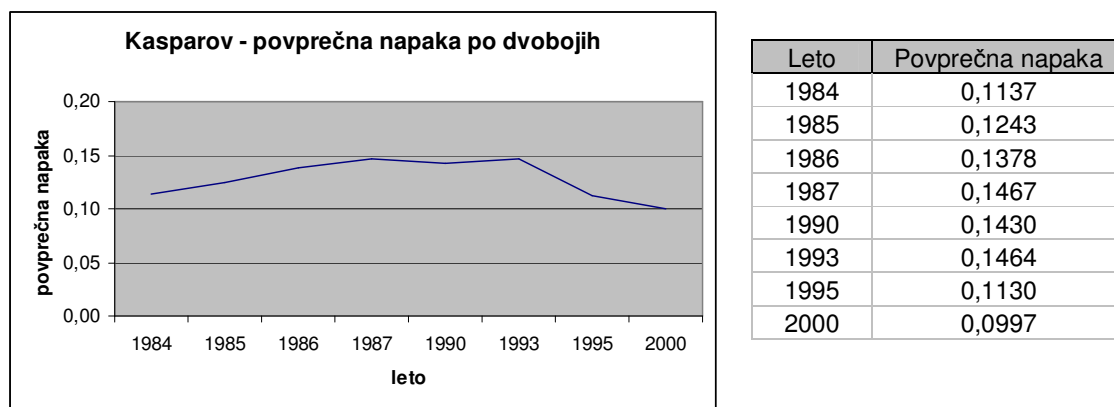
Leto	Povprečna napaka
1948	0,1709
1951	0,1552
1954	0,1290
1957	0,1982
1958	0,1676
1960	0,1865
1961	0,1528
1963	0,1183

Slika 3.3: Že na prvi pogled je očitno, da krivulja povprečnih napak pri Botvinniku poteka nekoliko višje kot pri Laskerju, Botvinnik je torej v povprečju delal večje napake. Najslabše v tem pogledu se je odrezal leta 1957 v dvoboju proti Smislovu (ko je tudi izgubil naslov), najmanjša povprečna napaka pa je zabeležena pri dvoboju s Petrosjanom leta 1963.

S primerjavo po dvobojih dobimo med drugim tudi vpogled, kateri dvoboji so najbolj vplivali na skupno oceno igralca. Slutiti nam da, da je vrednost napake lahko odvisna tudi od nasprotnika oz. njegovega stila igre. To je lepo vidno v primeru dvobojev Lasker-Steinitz

(1894 in 1896), ko je Steinitz s svojim bolj taktičnim načinom igre očitno vplival tudi na število napak nasprotnika. Nasprotno smo pri Botviniku najmanjšo povprečno napako zabeležili prav pri dvoboju z mirnim Petrosjanom.

Kot bomo videli v nadaljevanju, smo s pomočjo *kombiniranih kriterijev* vpliv nasprotnikov oz. različnih kompleksnosti pozicij, ki so nastajale v partijah različnih dvobojev, poskusili čimbolj izničiti s pomočjo simuliranja povprečnih napak pri povprečno kompleksnih partijah.



Slika 3.4: Pri Kasparovu ni opaziti takšnih nihanj v igri kot pri Botviniku. Ironija je, da je najbolje odigral dvoboj proti Kramniku leta 2000, katerega je edinega izgubil.

Analize povprečnih napak po dvobojih ostalih igralcev podajam v *Dodatku B*.

3.1.3 Primerjava povprečnih napak pri partijah posameznih dvobojev

Izvedel sem tudi nekatere primerjave povprečnih napak med igralci pri partijah posameznih dvobojev. Na ta način lahko npr. preverimo *formo* igralcev v posameznih stadijih dvoboja.

Eden izmed zanimivih dvobojev za tovrstno analizo je najdaljši dvoboj v zgodovini dvobojev za svetovnega prvaka, dvoboj med tedanjim svetovnim prvakom Karpovom in njegovim izzivalcem Kasparovom, ki je potekal v Moskvi od septembra 2004 do februarja 2005 in je trajal kar 48 partij!

Zmagovalec v tem dvoboju naj bi bil tisti, ki bi prvi dosegel šest zmag, remiji pa v tem dvoboju niso šteli k rezultatu. Karpov, ki je vodil že 5-0, nikakor ni uspel doseči še šeste zmage in po petih mesecih (!) borbe je bil dvoboj pri rezultatu 5-3 za Karpova zaradi njegovih zdravstvenih težav prekinjen.



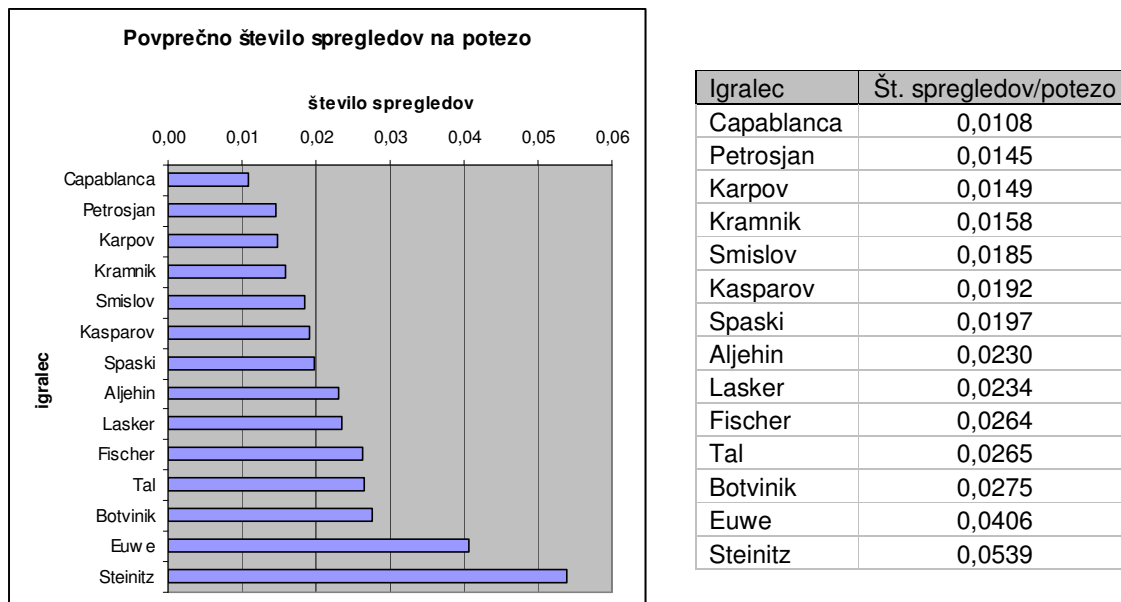
Slika 3.5: Grafična uprizoritev povprečnih vrednosti napak v posameznih partijah najdaljšega dvoboja za svetovnega prvaka v zgodovini

Iz grafa se nazorno vidi, da je Karpov (rdeča krivulja) v zadnjem delu dvoboja delal večje napake kot poprej, kar je verjetno botrovalo njegovi želji po prekinitvi dvoboja.

Odvisnost med odstopanji povprečnih napak obeh igralcev v posamezni partiji in rezultatom te partije sem pojasnil v poglavju 2.10.

3.1.4 Število grobih napak (spregledov)

Za grobo napako oz. spregled štejejo napake, ki presegajo numerično vrednost 1.00 oz. enega izgubljenega kmeta. Ugotavljanje povprečnega števila spregledov na potezo je bil eden izmed osnovnih kriterijev za ugotavljanje kvalitete igre igralcev. Čim manjše je to število pri nekem igralcu, tem boljši je ta igralec po tem kriteriju.



Slika 3.6: Povprečno število spregledov na potezo pri posameznih igralcih

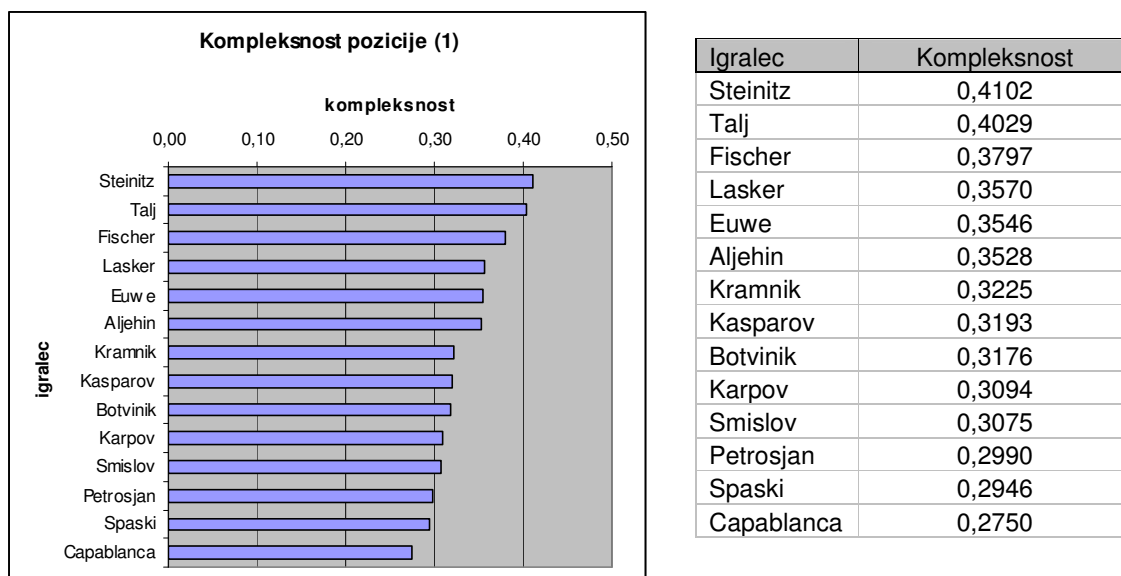
Tudi pri tem kriteriju je nesporni zmagovalec Capablanca, medtem ko je Steinitz prepričljivo zadnji. Lestvica je tudi sicer podobna tisti pri ugotovljenih povprečnih napakah igralcev. V primerjavi z njo izstopa Petrosjan, ki je tu na visokem drugem mestu. To ni presenetljivo, saj Petrosjan podobno kot Capablanca velja za izredno mirnega pozicijskega igralca, ki se izogiba taktičnim zapletom. V le-teh pa je možnost spregleda navadno veliko večja kot v mirnih pozicijah.

Povprečno število pri partijah vseh igralcev skupaj znaša 0,0239.

3.2 Težavnost odločanja (kompleksnost pozicije)

Rezultati ugotavljanja težavnosti odločanja oz. kompleksnosti pozicij v partijah obravnavanih igralcev so bili še posebej težko pričakovani, saj smo si želeli preveriti pravilnost naših hipotez oz. kriterijev za njegovo preverjanje.

Pričakovati je bilo nižja ugotovljena kompleksnost pozicij pozicijskih igralcev in višja pri taktičnih igralcih.



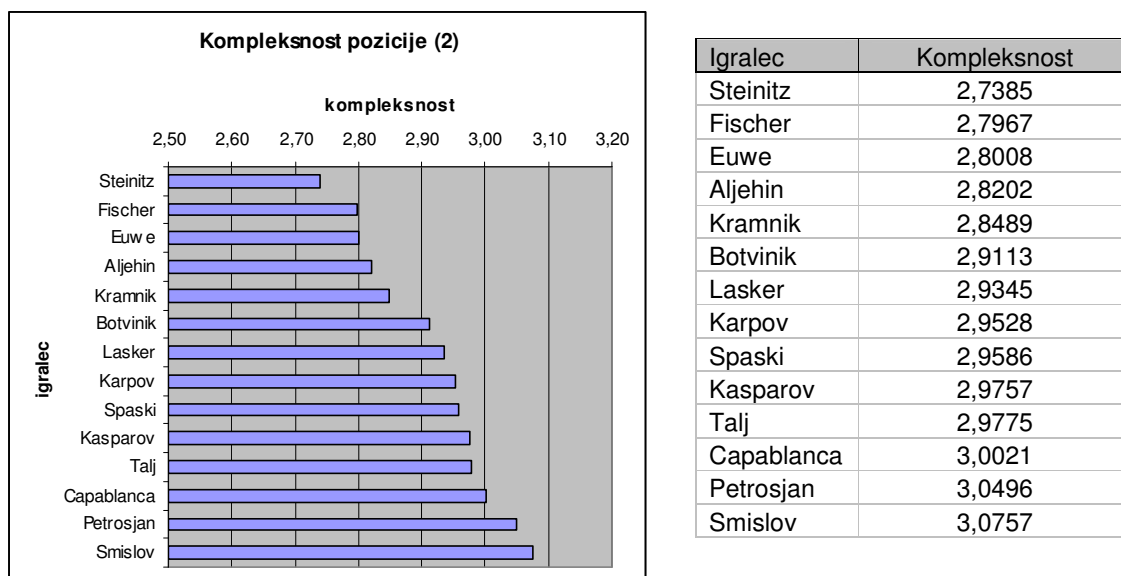
Slika 3.7: Povprečna kompleksnost pozicije pri posameznih igralcih (metoda 1)

Rezultati so povsem potrdili našo hipotezo, hkrati pa tudi potrdili korektnost uporabljene metode (kar sem podrobneje utemeljil v poglavju 2.10). Na vrhu lestvice po padajoči kompleksnosti sta tako najbolj izrazito taktična igralca Steinitz in Talj, medtem ko so z najmanj kompleksnimi pozicijami imeli opravka Capablanca, Spaski in Petrosjan, ki veljajo za tipične pozicijske igralce.

Izračunana povprečna kompleksnost pozicije v partijah vseh igralcev znaša 0,3353.

Zgornji rezultati so bili pridobljeni z metodo, pri kateri sem ugotavljal *vsoto* razlik med ocenami najboljših in drugih najboljših potez pri globinah iskanja v primerih, ko si je računalnik pri različnih globinah iskanja premislil, katera je najboljša poteza. To metodo smo opredelili kot glavno za preverjanje težavnosti odločanja pri igralcih.

Želeli pa smo preveriti še drugo metodo, pri kateri smo zabeležili *število* dogodkov, ko si je računalnik pri povečani globini premislil, katera je najboljša poteza. Rezultate, dobljene po tej metodi, prikazuje naslednja slika s priloženo tabelo.



Slika 3.8: Povprečna kompleksnost pozicije pri posameznih igralcih (metoda 2)

Izračunana povprečna kompleksnost pozicije v partijah vseh igralcev po tej metodi znaša 2,9135.

Podobnost lestvic, pridobljenih s pomočjo obeh metod za ocenjevanje kompleksnosti, je očitna že na prvi pogled. S to razliko, da je pri drugi metodi merilo za večjo kompleksnost čim *manjša* vrednost oz. število dogodkov, ko si je računalnik pri različnih globinah iskanja premislil glede najboljše poteze. To ni nelogično, saj je večje število preišljanj navadno značilno za pozicije, v katerih je na voljo več približno enakovrednih možnosti za izbiro najboljše poteze. V takih pozicijah je odločanje lažje in so po naših merilih manj kompleksne.

Po drugi strani so pozicije, kjer si računalnik pri različnih globinah iskanja sploh ne premisli glede opredelitve najboljše poteze, enostavne za odločanje, saj igralec v pomanjkanju alternativ navadno nima težke izbire.

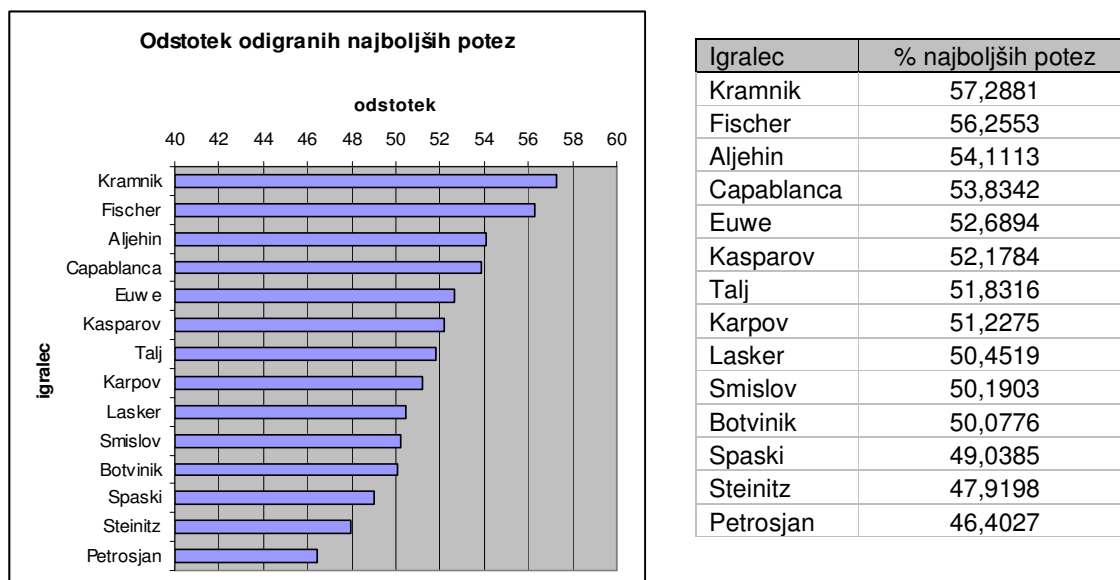
Dvom o korektnosti druge metode še dodatno podkrepi razlika v uvrstitvi Talja na obeh lestvicah. Po rezultatih, pridobljenih s pomočjo druge metode, so njegove pozicije bile približno tako (ne)kompleksne kakor pozicije od Capablance, za kar pa šahisti vemo, da ne drži.

Kot korektno metodo za ocenjevanje kompleksnosti smo torej na osnovi rezultatov opredelili prvo metodo, njeno korektnost pa sem še dodatno utemeljil z raziskavami, katerih rezultate sem prikazal v poglavju 2.10.

Metoda utegne biti izredno zanimiva za nadaljnje raziskovanje, saj se lahko izkaže za primerno tudi v drugih sistemih, kjer si pri odločanju pomagamo z odločitvenimi drevesi na podoben način, kot to počnejo računalniški šahovski programi. Ugotavljanje kompleksnosti odločanja pri tovrstnih sistemih bi utegnilo imeti veliko uporabno vrednost, saj bi med drugim lahko tudi pomenilo informacijo glede zanesljivosti pridobljenih ocen oz. odločitev. Tako bi se na osnovi pridobljene ocene glede kompleksnosti odločanja lahko npr. odločili za povečano globino iskanja itd.

3.3 Odstotek odigranih najboljših potez

Odstotek odigranih najboljših potez, kot se že pojasnil pri *Metodah dela*, sam po sebi ne opredeljuje neposredno kvalitete igralca, kot bi to morda kdo pričakoval. V nekaterih tipih pozicij je namreč občutno lažje najti potezo, ki odstopa od ostalih v pozitivnem smislu in je torej boljša od ostalih.

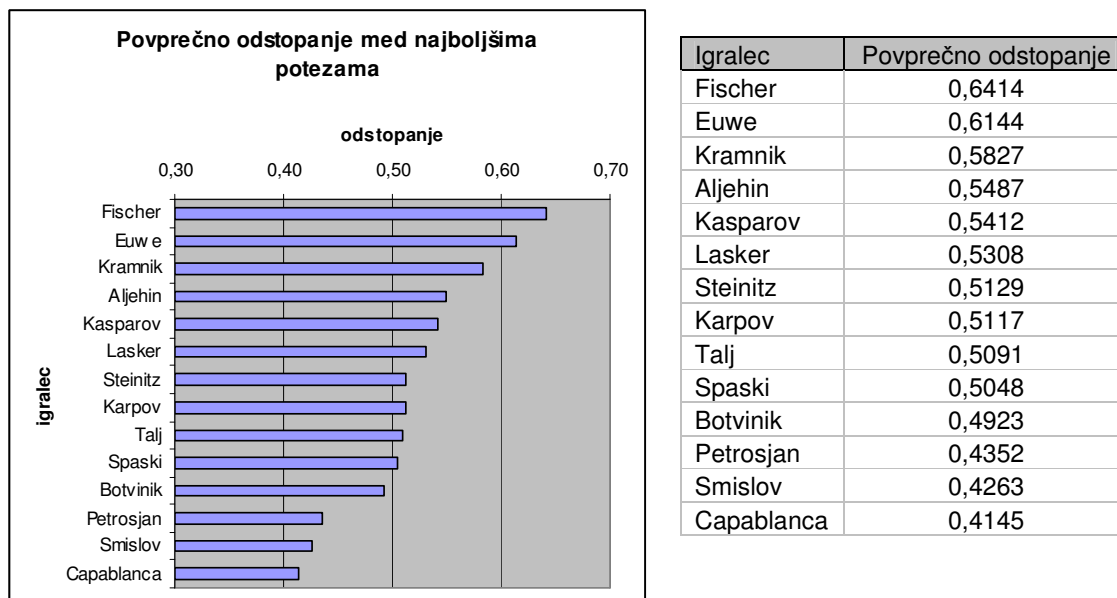


Slika 3.9: Odstotek odigranih najboljših potez pri posameznih igralcih

Izračunani povprečni odstotek odigranih najboljših potez v partijah vseh igralcev po tej metodi znaša 51,2378.

Potrebno je bilo torej ugotoviti, kaj determinira težavnost izbire najboljše ocene. Kot možen kriterij smo opredelili razliko med ocenama najboljših dveh potez v dani poziciji. Če je ta razlika večja, potem naj bi igralec imel olajšano izbiro.

Rezultati, pridobljeni po tem kriteriju, so prikazani na naslednji strani.



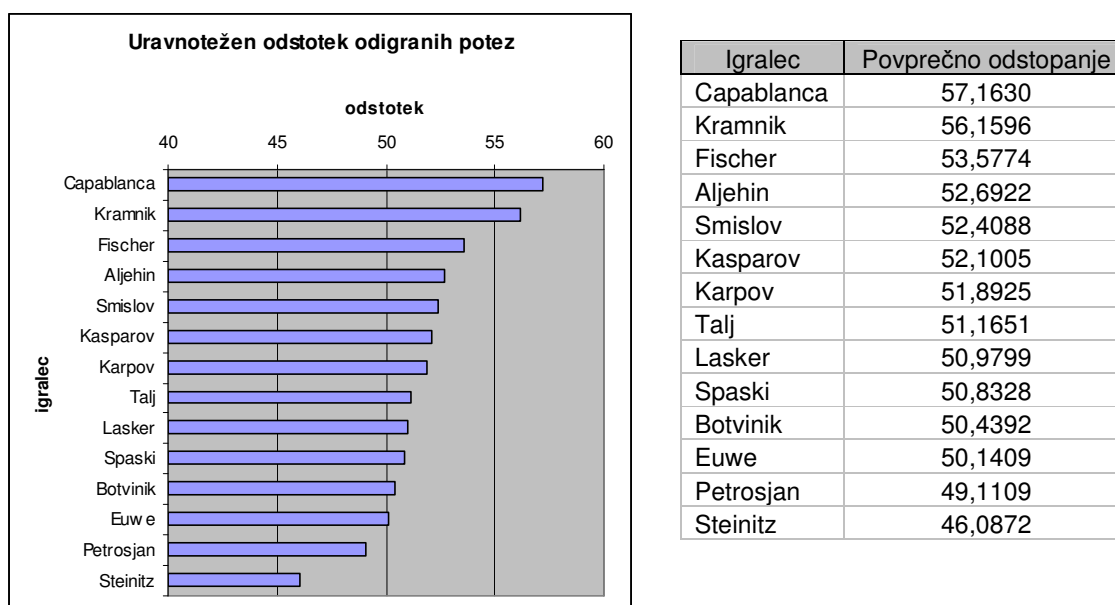
Slika 3.10: Povprečno odstopanje med najboljšima potezama pri posameznih igralcih

Povprečno odstopanje med najbolje ocenjenima potezama v partijah vseh igralcev je 0,5169.

Že na prvi pogled je mogoče potegniti vzporednico med rezultati obeh kriterijev. Lestvici sta približno enaki, pri vrhu obeh so Fischer, Kramnik in Aljehin, pri dnu pa Petrosjan, Spaski in Smislov. Lahko pa opazimo izrazita odstopanja nekaterih igralcev. Capablanca, ki je imel v povprečju najmanjša odstopanja, je na lestvici, ki prikazuje odstotek odigranih potez povsem pri vrhu. Medtem ko za Steinitza, ki je na lestvici odstopanj relativno visoko, velja ravno obratno: na prvi lestvici je povsem pri dnu. In to sta ravno igralca, ki sta po kriteriju izračunanih povprečnih napak na prvem in zadnjem mestu!

Iz tega je bilo moč sklepati, da bi se s primerno združitvijo obeh kriterijev dalo sklepati o kvaliteti igralcev. Capablanca npr. je imel po naši hipotezi najtežjo izbiro glede najboljših potez, kljub temu pa je odigral relativno visok odstotek najboljših potez, kar naj bi odražalo njegovo kvaliteto.

Tako sem izoblikoval *kombiniran kriterij* (podrobneje je predstavljen v *Metodah dela*), s pomočjo katerega smo simulirali odstotek najboljših odigranih potez pri povprečnem odstopanju med najboljšima potezama. Rezultati so predstavljeni na naslednji strani.



Slika 3.11 Simuliran odstotek odigranih najboljših potez, ki bi nastopil v primeru, da bi pri vseh igralcih bila enaka odstopanja med najboljšima potezama

Rezultati naj bi odražali odstotek najboljših potez, ki bi jih igralci povlekli, če bi vsi imeli opravka z enakimi pozicijami glede razlike med prvo in drugo najbolje ocenjeno potezo.

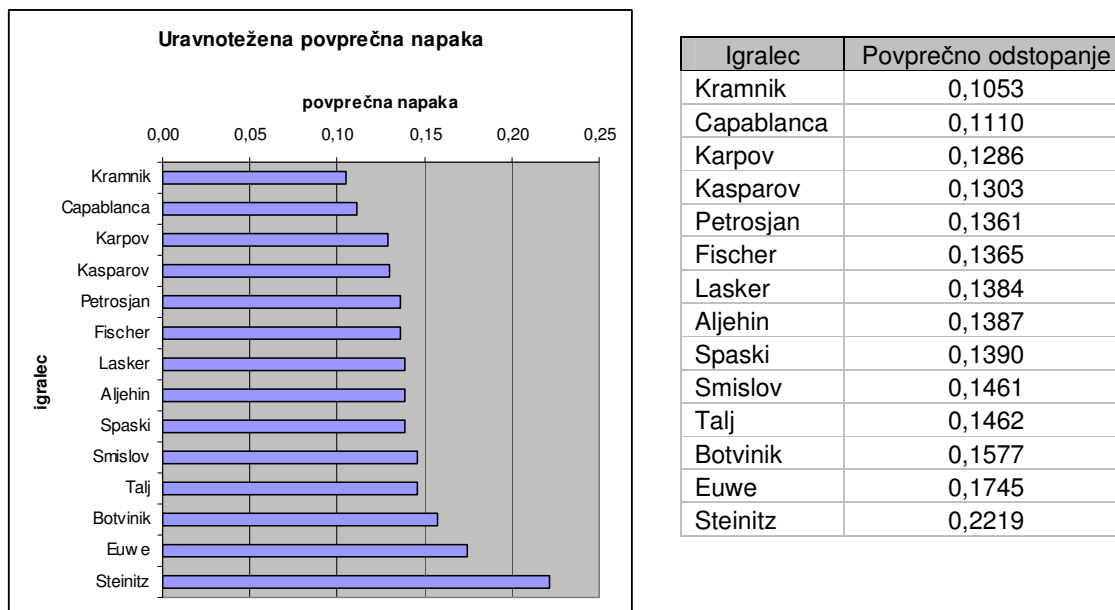
Tudi po tem kriteriju za ugotavljanje kvalitete igralcev je Capablanca zmagovalec, Steinitz pa povsem na dnu. V primerjavi s kriterijem, kjer smo ugotavljali povprečno napako igralcev, lahko opazimo še zelo podobne uvrstitve Kramnika, Botvinika in Euweja, medtem ko sta najbolj napredovala taktična igralca Fischer in Aljehin, najbolj pa je nazadoval pozicijski igralec Petrosjan (ki je navsezadnje tudi povlekel najmanjši odstotek najboljših potez).

Rezultat je torej logičen, saj je kriterij *povprečna napaka* nekoliko postavljaj v prednost pozicijske igralce.

3.4 Napaka igralcev pri uravnoteženi kompleksnosti pozicij

Z namenom, da bi odgovoril na vprašanje, kdo je bil najboljši svetovni prvak vseh časov, vsaj kar se tiče partij, odigranih v dvobojih za naslov svetovnega prvaka, sem oblikoval še en kombiniran kriterij, katerega namen je bil ugotoviti povprečne napake igralcev, če bi le-ti igrali enako kompleksne pozicije.

Slabost kriterija ugotavljanja povprečnih napak igralcev je bila namreč ta, da je nekoliko favoriziral pozicijske igralce, saj so le-ti v manj kompleksnih pozicijah, s katerimi so navadno imeli opravka, delali manj napak.



Slika 3.12 Simulirana povprečna napaka, ki bi nastopila v primeru, da bi vsi igralci imeli opravka z enako kompleksnimi pozicijami (vzeto je povprečje kompleksnosti pozicij vseh igralcev)

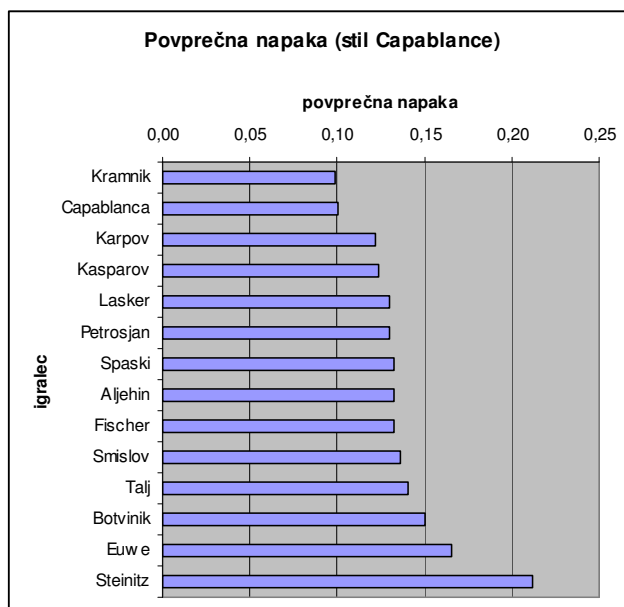
Tako sem izvedel simulacijo, s pomočjo katere sem lahko ugotovil pričakovane vrednosti napak, ki bi jih igralci naredili v partiji, ki bi predstavljala povprečno kompleksno partijo

glede na vse njihove partije. Podrobnosti tega kriterija sem natančneje opisal v *Metodah dela*, rezultati pa so prikazani na naslednji sliki s priloženo tabelo.

Zamenjava Kramnika in Capablance na prvih dveh mestih v primerjavi z rezultati kriterija *povprečna napaka* je bila pričakovana, saj sta oba v povprečju delala približno enako velike napake, medtem ko je Kramnik igral nekoliko kompleksnejše pozicije. Pri svojih ocenah so v primerjavi z omenjenim kriterijem napredovali vsi, ki so imeli opravka z nadpovprečno kompleksnimi pozicijami, nazadovali pa vsi, ki so igrali enostavnejše pozicije glede na povprečje vseh igralcev.

Pri tem kombiniranem kriteriju ne smemo pozabiti, da smo za povprečje vzeli povprečno kompleksno partijo tu obravnavanih igralcev in ne povprečno kompleksne partije na splošno.

Kakšni bi bili rezultati, če bi vsi igralci igrali tako kompleksne pozicije, kot jih je igral Capablance? Odgovor na to vprašanje skušam podati na naslednji sliki s priloženo tabelo, kjer so prikazani rezultati, ki jih dobimo, če bi za povprečno kompleksno partijo vzeli povprečno kompleksno partijo Capablance.

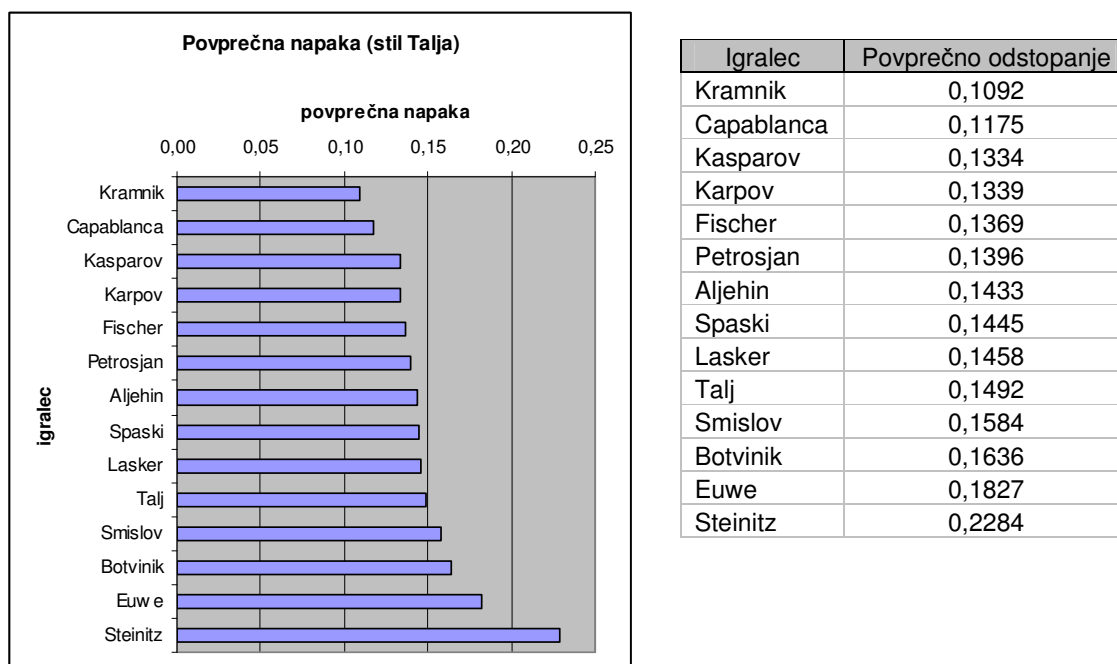


Igralec	Povprečno odstopanje
Kramnik	0,0987
Capablance	0,1008
Karpov	0,1214
Kasparov	0,1240
Lasker	0,1295
Petrosjan	0,1303
Spaski	0,1322
Aljehin	0,1325
Fischer	0,1330
Smislov	0,1360
Talj	0,1403
Botvinik	0,1507
Euwe	0,1655
Steinitz	0,2114

Slika 3.13 Simulirana povprečna napaka, ki bi nastopila v primeru, da bi vsi igralci imeli opravka z enako kompleksnimi pozicijami, kot jih je igral Capablance

Ker je Capablanca igral najmanj kompleksne pozicije med vsemi igralci, se je vsem igralcem nekoliko poboljšal rezultat (v povprečju bi delali manjše napake), le njegov rezultat ostaja enak. Spet sta Kramnik in Capablanca prepričljivo na prvih mestih.

Zanimalo me je tudi, kakšni bi bili rezultati, če bi vsi igrali v Taljevem stilu oz. imeli opravka z enako kompleksnimi pozicijami kot on.



Slika 3.14 Simulirana povprečna napaka, ki bi nastopila v primeru, da bi vsi igralci imeli opravka z enako kompleksnimi pozicijami, kot jih je igral Talj

Tudi pri tako kompleksnih pozicijah, kot jih je imel Talj, bi Kramnik v povprečju delal najmanjše napake. Rezultati vseh igralcev so nekoliko slabši kot pri rezultatih osnovnega kriterija *povprečna napaka*, Taljev rezultat je enak, izboljšal se je edino rezultat Steinitza, ki je imel opravka še z nekoliko kompleksnejšimi pozicijami kot Talj.

3.5 Material tekom partije

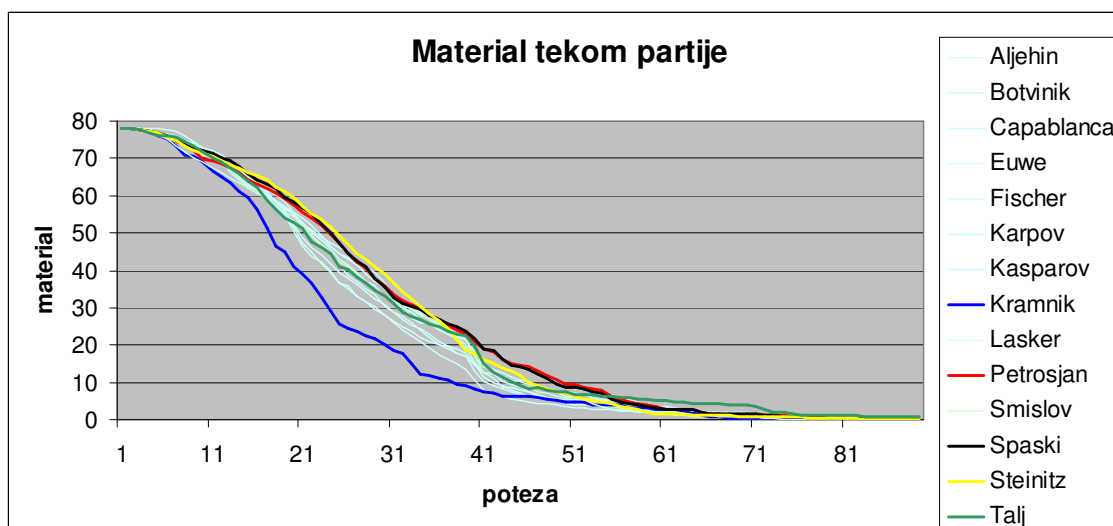
Namen ugotavljanja povprečne količine materiala na šahovnici tekom partije ni bil preverjati kvalitete igre obravnavanih igralcev, pač pa dobiti dodatne informacije o načinu njihove igre. Predvsem sem želel preveriti njihove *težnje po poenostavitvah pozicije*.

Vrednost materiala je izražena numerično, posamezne vrednosti figur pa so predstavljene na naslednji tabeli:

Figura	Vrednost
kralj	-
dama	9
trdnjava	5
lovec	3
skakač	3
kmet	1

Tabela 3.2: Numerične vrednosti posameznih figur

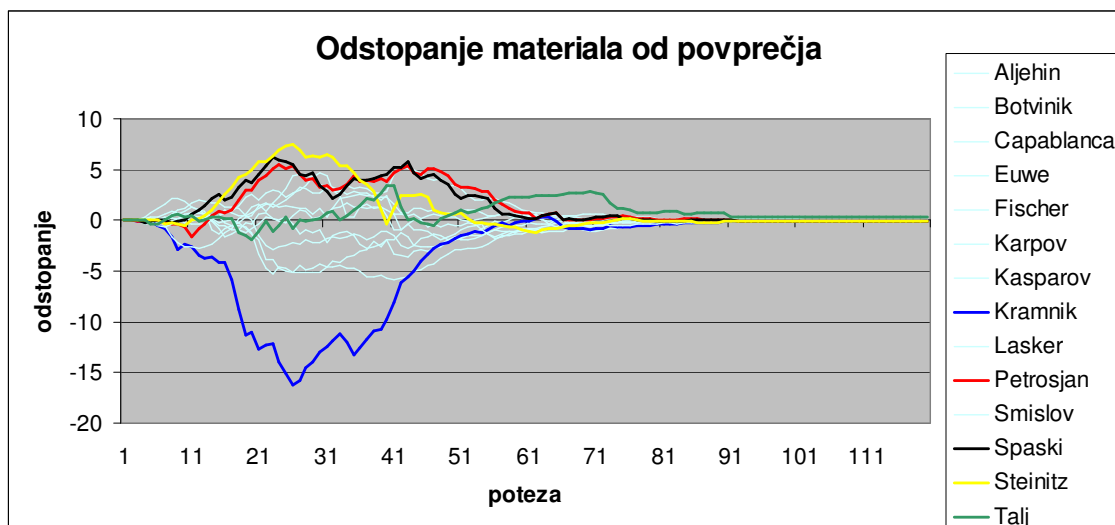
Vrednost materiala na šahovnici v začetni poziciji znaša 39 za bele in 39 za črne, skupaj torej 78. Naslednja slika nazorno prikazuje padanje vrednosti materiala tekom partije pri posameznih igralcih.



Slika 3.15: Material tekom partije

Poudarjeni so igralci, ki izstopajo glede menjav figur. To so Kramnik (modra krivulja), Petrosjan (rdeča), Spaski (črna), Steinitz (rumena) in Talj (zelena).

Odstopanja med posameznimi igralci glede količine materiala na šahovnici pri posameznih potezah prikazuje naslednja slika. Za povprečje je vzet povprečen material v vseh njihovih partijah.



Slika 3.16: Odstopanje materiala od povprečja

Že na prvi pogled je očitno izrazito odstopanje Kramnika, ki je imel z naskokom najmanjšo količino materiala na šahovnici med 20. in 40. potezo, kar je gotovo tudi posledica zgodnjih menjav dam s črnimi figurami v dvoboju proti Kasparovu.

Po nekoliko večji količini materiala na šahovnici odstopajo Steinitz, Petrosjan in Spaski, zanimiva pa je še krivulja Talja, ki je očitno odigral kar nekaj partij, v katerih je še v zadnjem stadiju partije bilo na šahovnici precej figur.

H končnicam je torej najbolj stremel Kramnik, medtem ko bi predvsem za Steinitza lahko trdili ravno obratno.

4

Sklep

Izvesti računalniško primerjavo svetovnih šahovskih prvakov ni bil edini namen tega dela. Že pred začetkom so se namreč pojavila vprašanja, na katera ni bilo mogoče odgovoriti brez eksperimentalnega dela. So računalniški šahovski programi že dovolj dobri, da bi lahko na osebnih računalnikih uspešno izvedli primerjavo najboljših igralcev v zgodovini? Bo mogoče katerega od šahovskih programov odprte kode spremeniti tako, da bo v primernem času mogoče izvesti dovolj veliko število analiz partij? Namen je bil tudi poiskati čim več uporabnih kriterijev za ocenjevanje in izoblikovati metode dela, ki bodo lahko služile tudi za podobne raziskave v prihodnosti, z močnejšimi računalniki in boljšimi šahovskimi programi.

Pri svojem delu sem se soočil tudi s pojmom *kompleksnost pozicije in stil igranja*. Poskusil sem ugotoviti pravo metodo za določitev, kako kompleksna oz. težavna za odločanje je posamezna pozicija, ter dognati, kaj določa stil posameznih igralcev. Oboje mi je pomagalo pri izoblikovanju *kombiniranih kriterijev*, s katerimi sem želel dobiti še nekoliko objektivnejše rezultate glede kvalitete posameznih igralcev.

Izkazalo se je, da na zgornji dve vprašanji lahko odgovorimo pritrdilno. Crafty, najmočnejši program odprte kode na svetu, se je izkazal kot primerna izbira orodja, saj se ga je dalo preprogramirati tako, da je ustrezal našim namenom, izkazal pa se je tudi kot verodostojen ocenjevalec.

Rezultati so potrdili našo hipotezo, da so mirni pozicijski igralci, ki so igrali manj kompleksne pozicije, delali manj napak v igri kot ostri taktični igralci. Med pozicijske igralce se uvršča tudi tretji svetovni prvak Jose Raul Capablanca, ki je nesporni zmagovalec po večini kriterijev, ki so bili namenjeni ocenjevanju kvalitete igre. Capablanca se je izkazal za najbolj natančnega med vsemi svetovnimi prvaki, v njegovi igri smo zabeležili najmanj napak.

Ali to pomeni, da je bil Capablanca najboljši šahist med vsemi svetovnimi šahovskimi prvaki?

Rezultati so nazorno pokazali, da je število napak odvisno tudi od kompleksnosti pozicij, s katerimi so se šahisti soočali. Capablanca pa je med vsemi prvaki imel opravka z najmanj kompleksnimi pozicijami. Da bi ugotovili, kako bi se igralci odrezali v enako kompleksnih pozicijah, sem izvedel raziskave, namen katerih je bil dognati pričakovano število njihovih napak v tem primeru. Zmagovalec pri teh raziskavah je najmlajši svetovni prvak Vladimir Kramnik, ki bi po dobljenih rezultatih delal najmanjše napake, tudi če bi igral v stilu Capablance. Capablanca je bil pri teh raziskavah še vedno na visokem drugem mestu. V njegov prid lahko navedemo, da so igralci pri izbiri stila imeli proste roke. Če so uspeli tudi nasprotnikom vsiliti stil igre, v katerem so bili sami najbolj uspešni, je bil tudi to navadno odraz njihove kvalitete.

Odločitev, kdo je bil najboljši, prepuščam Vam, vsekakor pa gre izbirati med Capablancem in Kramnikom, ki po rezultatih prepričljivo odstopata od ostalih, seveda v pozitivnem smislu.

Pri svojem delu sem imel še največ težav z odkrivanjem morebitnih napak pri pridobivanju rezultatov, saj je bilo zaradi ogromne količine vhodnih podatkov to delo močno oteženo. V ta namen sem napisal številne skripte za testiranje in primerjavo rezultatov, vse rezultate sem tudi sproti preverjal, pri tem pa sem si v dobršni meri pomagal tudi s šahovskim znanjem. Z veliko mero gotovosti lahko zatrdim, da tudi če so se kljub vsemu v obdelane podatke slučajno prikradle kakšne drobne napake, le-te niso imele večjega vpliva na končne rezultate.

Natančnost ocen bi se dalo izboljšati z ocenjevanjem še večjega števila partij, povečano globino iskanja in z uporabo močnejših šahovskih programov, potem ko bodo ti na voljo. Za ocenjevanje bi se dalo uporabiti tudi več računalniških šahovskih programov in vzeti za merodajne povprečja njihovih ocen pri posameznih potezah.

Nadaljnje delo na tem področju bi lahko bilo usmerjeno v oblikovanje še večjega števila kriterijev za ocenjevanje ter v morebitno izboljšanje že obstoječih. Še posebej zanimivo v tem pogledu je ugotavljanje težavnosti odločanja pri preiskovanju iskalnih dreves, ki utegne biti zanimivo tudi za širše področje umetne inteligence.

Zahvala

Zahvaljujem se mentorju prof. dr. Ivanu Bratku za čudovito temo, ki sem se je z veseljem lotil, in za veliko mero strokovne pomoči, asistentu mag. Aleksandru Sadikovu za številne uporabne napotke ter Gorazdu Lampiču za pomoč pri nekaterih kompleksnejših izračunih. Hvala tudi dr. Janezu Demšarju in Martinu Možini za napotke glede izračunov statističnih podatkov ter vele mojstru Draganu Šolaku in mednarodnemu mojstru Marku Tratarju za pomoč pri odločanju pri nekaterih šahovskih vprašanjih. Pa tudi družini in vsem ostalim bližnjim za izkazano moralno podporo. Brez njih vseh to delo ne bi bilo tako, kot je. Hvala.

Literatura

- [1] Bratko, I. (2000). *Prolog Programming for Artificial Intelligence*. Addison-Wesley, Reading, Massachusetts, tretja izdaja.
- [2] Russell, S. in Norvig, P. (2002). Adversarial Search. V *Artificial Intelligence: A Modern Approach*. Prentice Hall, druga izdaja.
- [3] Kasparov, G. in Plisecki, D. (2004). *Moji veliki predhodniki*. Šahohlačnik.
- [4] *Game Theory*: <http://www.maths.nott.ac.uk/personal/anw/G13GT1/compch.html>.
- [5] *Chess Tree Search*: <http://chess.verhelst.org/1997/03/10/search/>.
- [6] *Dr. Robert Hyatt's home page*: <http://www.cis.uab.edu/info/faculty/hyatt/hyatt.html>.
- [7] *Svenska schackdatorföreningen (SSDF)*: <http://web.telia.com/~u85924109/ssdf/>.
- [8] *The Crafty ftp site*: <ftp://ftp.cis.uab.edu/pub/hyatt/>.
- [9] Luštrek, M. (2002). *Računalniško igranje iger s kartami*. Odsek za inteligentne sisteme, Institut Jožef Stefan.
- [10] Sadikov, A. (2002). *Kako se računalnik spopade z igrami?* Laboratorij za Umetno Inteligenco, Fakulteta za računalništvo in informatiko.
- [11] *World Chess Championships*: <http://chess.about.com/od/worldchampionship/>.
- [12] *Chessmetrics*: <http://www.chessmetrics.com>

Dodatek

A Izvorna koda spremenjenega Craftyja

Crafty obsega več kot 70 datotek z izvorno kodo in za dosego zastavljenih ciljev, še posebej pa za čim večjo uporabnost izpisov analiz, je bilo potrebno kar precej sprememb. Tukaj navajam le najbolj bistvene, v pomoč programerjem, ki se bodo v prihodnje lotevali podobnih nalog.

Annotate.c

Ključna sprememba je pravilna umestitev spodnje *for* zanke na pravo mesto v kodo. Zanka poskrbi za izvajanje analiz pri različnih fiksnih globinah, do maksimalne fiksne globine. Pred njo podajam še par vrstic kode, ki poskrbijo za izpis diagrama in vrednosti materiala v izhodno html datoteko ter za izpise tudi pri nižji globinah iskanja.

```

240:  }
241:  analysis_printed=0;

// izpiši diagram
if (strlen(html_br)) AnnotatePositionHTML(tree,wtm,annotate_out);

// izpiši podatke o materialu
fprintf(annotate_out,"white:%d black:%d total:%d%s%s\n",
        tree->pos.white_pieces+tree->pos.white_pawns, tree->pos.black_pieces +
        tree->pos.black_pawns, tree->pos.white_pieces+tree->pos.white_pawns +
        tree->pos.black_pieces+tree->pos.black_pawns,html_br,html_br);

// poskrbi za izpise tudi pri nižjih globinah
noise_level = 1;

// ponavljaj zanko pri različnih fiksnih globinah, do maksimalne fiksne globine
for(search_depth=1; search_depth<=max_fixed_depth; search_depth++) {

    /// shrani vrednost search_depth v začasne spremenljivke
    temp_search_depth = search_depth;
    mgcurrent_depth = search_depth;

242:  if (move_number >= line1 && move_number <= line2) {
243:  if (annotate_wtm==2 || annotate_wtm==wtm) {

```

Omenjena zanka se zaključi malo več kot sto vrstic kasneje.

```

350:         }
351:     }

    // preberi search_depth iz naše začasne spremenljivke
    search_depth = mgcurrent_depth;
} // zaključek for zanke

352: /*
353: -----

```

Še pred pričetkom izvajanja zanke definiramo vrednosti globalnih spremenljivk, ki smo jih v ta namen deklarirali v chess.h. `START_SEARCH_DEPTH` je konstanta z vrednostjo 12 in pomeni začetno maksimalno fiksno globino iskanja.

```

86: int save_swindle_mode, save_asymmetry;
87: int html_mode=0;

max_fixed_depth = START_SEARCH_DEPTH;
mgcurrent_depth = 0;

88: /*
89: -----

```

Izbris naslednje vrstice iz kode poskrbi za izpise tudi najboljših potez, ne le napak.

```

320: if (player_pv.pathd>1 && player_pv.pathl>=1 &&
321: player_score+annotate_margin<annotate_score[0] &&
322: (temp[0].path[1]!=player_pv.path[1] || annotate_margin<0.0 ||
323: best_moves!=1)) {

```

Data.c

S spremembo vrednosti spremenljivke `king_safety_asymmetry` iz -40 na 0 dosežemo, da Crafty v svojih ocenah ni več defenziven.

```

612: int king_safety_asymmetry = 0;

```

Iterate.c

Na naslednji način ob prehodu v končnico povečamo globino iskanja.

```
408:  if (iteration_depth>3 && value>MATE-300 &&
409:      value>=(MATE-iteration_depth-1) && value>last_mate_score) break;

// povečaj globino iskanja ob prehodu v končnico
if (TotalWhitePieces < 15 && TotalBlackPieces < 15) {
    max_fixed_depth = START_SEARCH_DEPTH + 1;
}
else {
    max_fixed_depth = START_SEARCH_DEPTH;
}

410:  if ((iteration_depth >= search_depth) && search_depth) break;
411:  if (time_abort || abort_search) break;
```

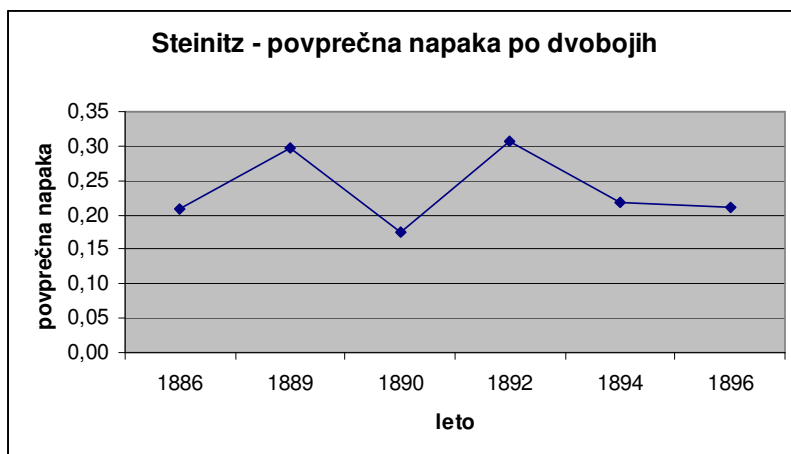
Ostale spremembe so v glavnem vezane na vsebino in obliko izpisov v izhodno datoteko.

Veliko zahtevnejši posegi so bili potrebni za korektno onemogočenje iskanja mirovanja, vendar ker so poskusi z onemogočeno tovrstno hevrstiko pokazali, da jo je za uresničitev našega namena neobhodno potrebno vključiti, tistih sprememb tukaj ne navajam.

B Analiza napak svetovnih prvakov po posameznih dvobojih

V nadaljevanju so prikazane povprečne napake svetovnih prvakov v posameznih dvobojih. Zaradi večje preglednosti so rezultati vsakega od prvakov prikazani na svoji strani.

I Wilhelm Steinitz

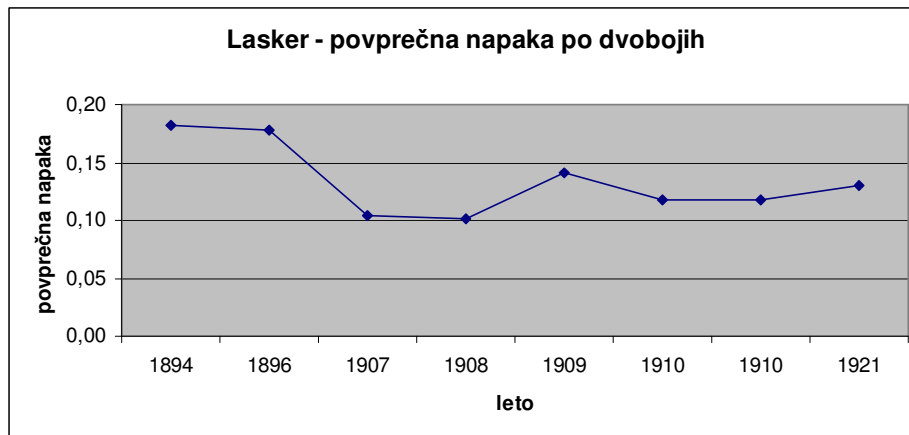


Slika B1: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Steinitzu

Leto	Napaka	Nasprotnik	Rezultat
1886	0,2087	Zukertort	12.5-7.5
1889	0,2965	Chigorin	10.5-6.5
1890	0,1740	Gunsberg	10.5-8.5
1892	0,3068	Chigorin	12.5-10.5
1894	0,2174	Lasker	7-12
1896	0,2098	Lasker	4.5-12.5

Tabela B1: Podatki o dvobojih

II Emanuel Lasker

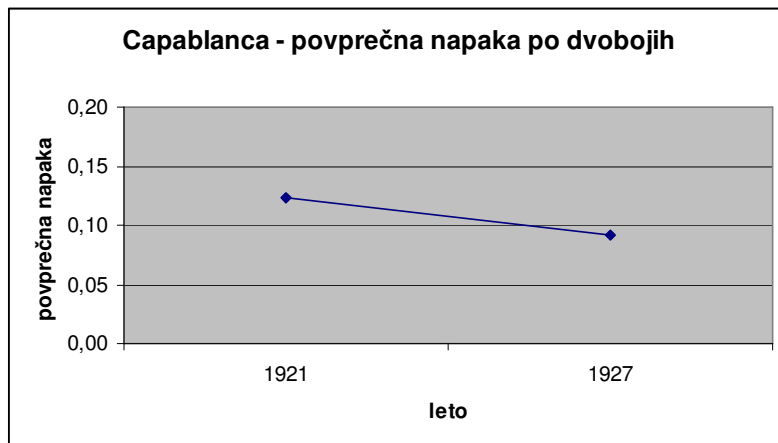


Slika B2: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Laskerju

Leto	Napaka	Nasprotnik	Rezultat
1894	0,1817	Steinitz	12-7
1896	0,1784	Steinitz	12.5-4.5
1907	0,1044	Marshall	11.5-3.5
1908	0,1007	Tarrasch	10.5-5.5
1909	0,1414	Janowski	8-2
1910	0,1173	Schlechter	5-5
1910	0,1184	Janowski	9.5-1.5
1921	0,1299	Capablanca	5-9

Tabela B2: Podatki o dvobojih

III Jose Raul Capablanca

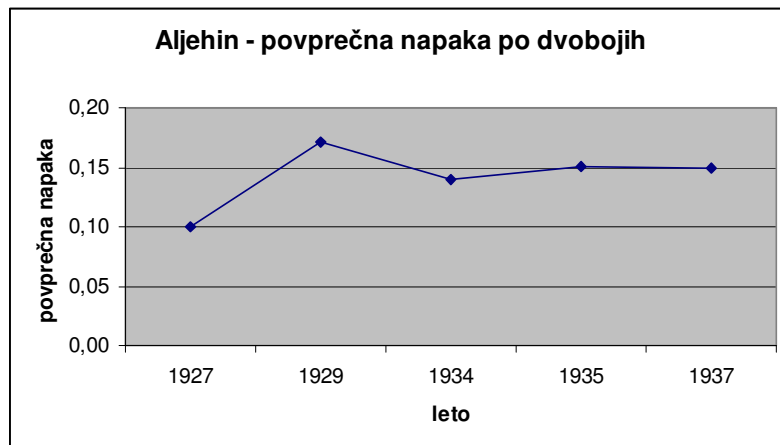


Slika B3: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Capablanci

Leto	Napaka	Nasprotnik	Rezultat
1921	0,1227	Lasker	9-5
1927	0,0922	Aljehin	15.5-18.5

Tabela B3: Podatki o dvobojih

IV Aleksander Aljehin

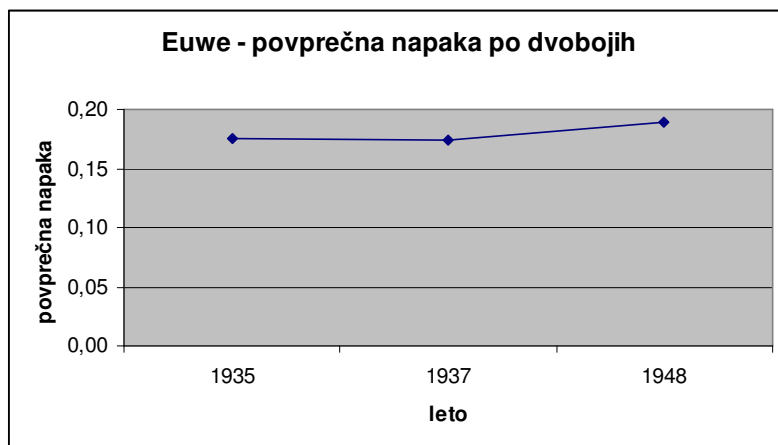


Slika B4: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Aljehinu

Leto	Napaka	Nasprotnik	Rezultat
1927	0,0997	Capablanca	18.5-15.5
1929	0,1719	Bogoljubov	15.5-9.5
1934	0,1403	Bogoljubov	15.5-10.5
1935	0,1502	Euwe	14.5-15.5
1937	0,1489	Euwe	17.5-12.5

Tabela B4: Podatki o dvobojih

V Maks Euwe

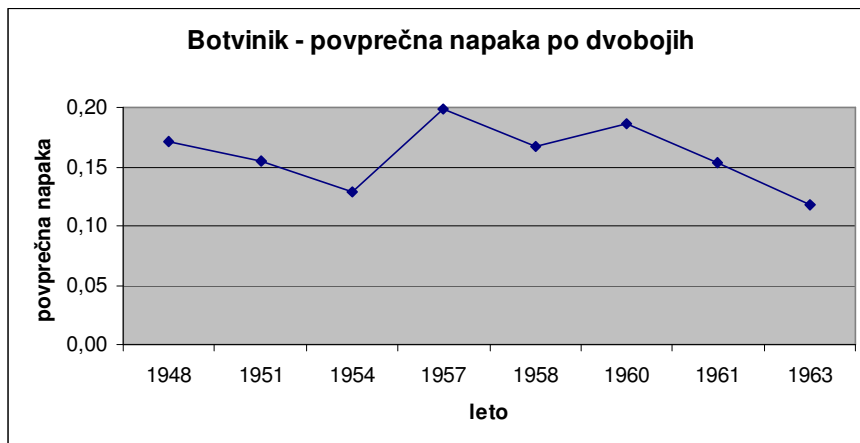


Slika B5: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Euweju

Leto	Napaka	Nasprotnik	Rezultat
1935	0,1750	Aljehin	15.5-14.5
1937	0,1740	Aljehin	12.5-17.5
1948	0,1889	(turnir)	4-16

Tabela B5: Podatki o dvobojih

VI Mihail Botvnik

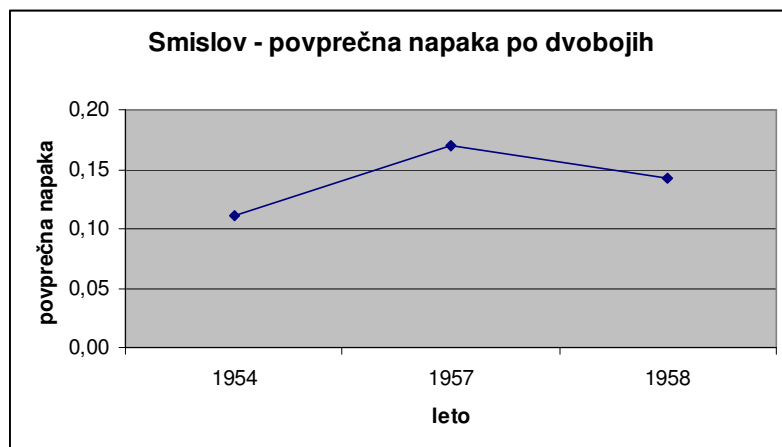


Slika B6: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Botvniku

Leto	Napaka	Nasprotnik	Rezultat
1948	0,1709	(turnir)	14-6
1951	0,1552	Bronstein	12-12
1954	0,1290	Smislov	12-12
1957	0,1982	Smislov	9.5-12.5
1958	0,1676	Smislov	12.5-10.5
1960	0,1865	Talj	8.5-12.5
1961	0,1528	Talj	13-8
1963	0,1183	Petrosjan	9.5-12.5

Tabela B6: Podatki o dvobojih

VII Vasilij Smislov

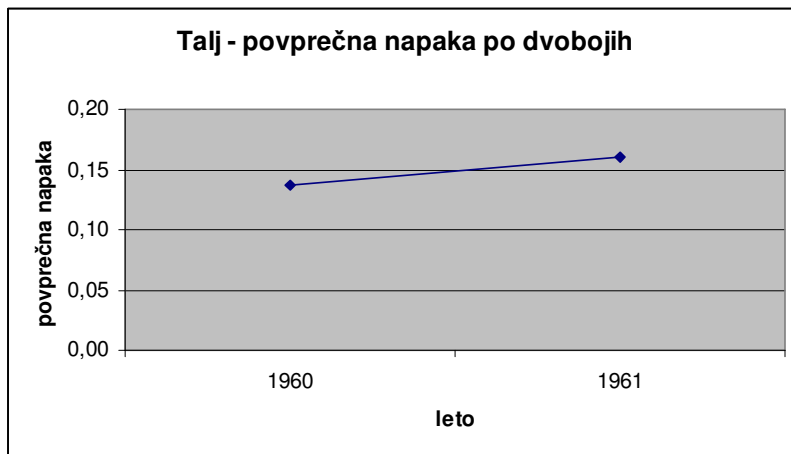


Slika B7: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Smislovu

Leto	Napaka	Nasprotnik	Rezultat
1954	0,1110	Botvinik	12-12
1957	0,1699	Botvinik	12.5-9.5
1958	0,1423	Botvinik	10.5-12.5

Tabela B7: Podatki o dvobojih

VIII Mihail Talj

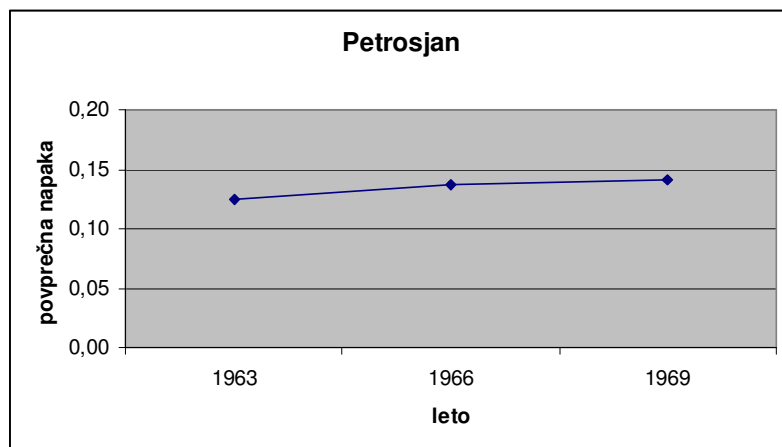


Slika B8: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Talju

Leto	Napaka	Nasprotnik	Rezultat
1960	0,1366	Botvinnik	12.5-8.5
1961	0,1597	Botvinnik	8-13

Tabela B8: Podatki o dvobojih

IX Tigran Petrosjan

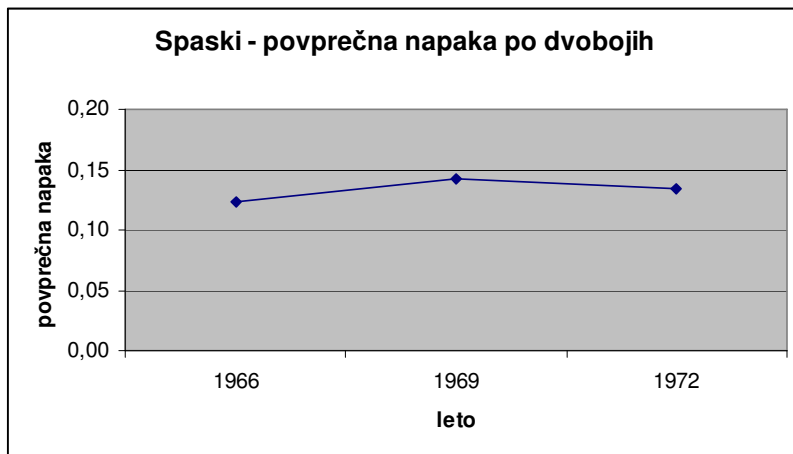


Slika B9: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Petrosjanu

Leto	Napaka	Nasprotnik	Rezultat
1963	0,1243	Botvinnik	12.5-9.5
1966	0,1371	Spaski	12.5-11.5
1969	0,1417	Spaski	10.5-12.5

Tabela B9: Podatki o dvobojih

X Boris Spaski

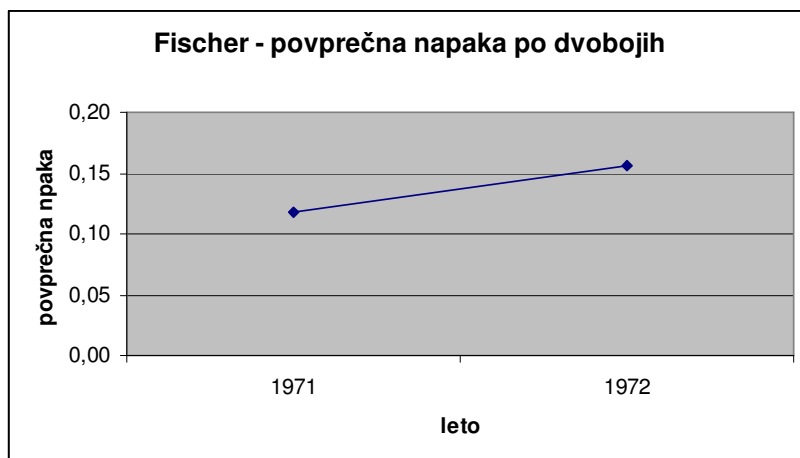


Slika B10: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Spaskem

Leto	Napaka	Nasprotnik	Rezultat
1966	0,1239	Petrosjan	11.5-12.5
1969	0,1427	Petrosjan	12.5-10.5
1972	0,1343	Fischer	7.5-12.5

Tabela B10: Podatki o dvobojih

XI Robert James Fischer

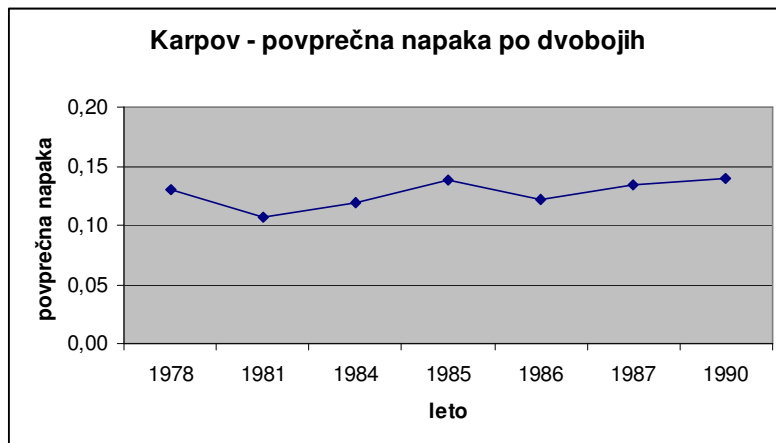


Slika B11: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Fischerju

Leto	Napaka	Nasprotnik	Rezultat
1971	0,1183	(kandidati)	18.5-2.5
1972	0,1568	Spaski	12.5-7.5

Tabela B11: Podatki o dvobojih

XII Anatolij Karpov

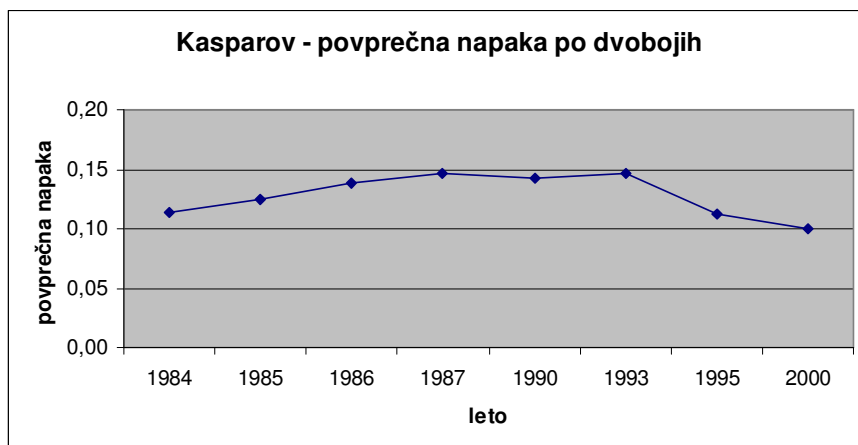


Slika B12: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Karpovu

Leto	Napaka	Nasprotnik	Rezultat
1978	0,13	Korčnoj	16.5-15.5
1981	0,1075	Korčnoj	11-7
1984	0,1193	Kasparov	25-23
1985	0,1384	Kasparov	11-13
1986	0,122	Kasparov	11.5-12.5
1987	0,1339	Kasparov	12-12
1990	0,1395	Kasparov	11.5-12.5

Tabela B12: Podatki o dvobojih

XIII Gari Kasparov

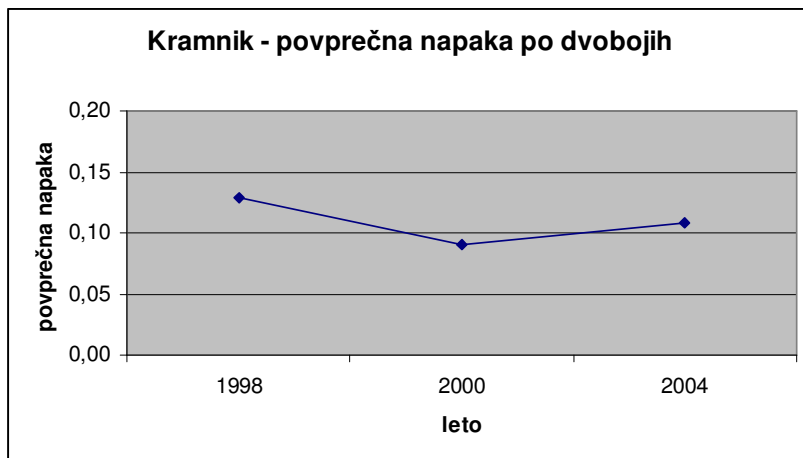


Slika B13: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Kasparovu

Leto	Napaka	Nasprotnik	Rezultat
1984	0,1137	Karpov	23-25
1985	0,1243	Karpov	13-11
1986	0,1378	Karpov	12.5-11.5
1987	0,1467	Karpov	12-12
1990	0,1430	Karpov	12.5-11.5
1993	0,1464	Short	12.5-7.5
1995	0,1130	Anand	10.5-7.5
2000	0,0997	Kramnik	6.5-8.5

Tabela B13: Podatki o dvobojih

XIV Vladimir Kramnik



Slika B14: Analiza povprečnih napak v dvobojih za svetovnega prvaka pri Kramniku

Leto	Napaka	Nasprotnik	Rezultat
1998	0,1290	Shirov	3.5-5.5
2000	0,0903	Kasparov	8.5-6.5
2004	0,1089	Leko	7-7

Tabela B14: Podatki o dvobojih

Pri dvoboju med Kasparovom in Kramnikom iz leta 2000 smo pri slednjem zabeležili v povprečju najmanjšo napako pri posameznih potezah od vseh dvobojev za svetovnega prvaka v zgodovini šaha.