

ACM MARIBOR

## **2. Delavnica sekcije ACM Maribor 2016**

Zbornik povzetkov



Urednika: Uroš Mlakar in Štefan Kohek

Maribor, 15. September 2016

# ACM MARIBOR

2. Delavnica sekcije ACM Maribor 2016

Maribor, 15. 9. 2016

Urednika: Uroš Mlakar in Štefan Kohek

Obdelava in oblikovanje: Uroš Mlakar in Štefan Kohek

Izdajatelj: ACM Maribor

Spletna izdaja: <http://acm.um.si/delavnica2016>

Leto izdaje: 2016

Organizacijski odbor:

Uroš Mlakar

Štefan Kohek

Iztok Fister Jr.

## Kazalo

Miha Ravber - Analiza indikatorjev uspešnosti v večkriterijski optimizaciji	4
Štefan Kohek - Prikaz senc pri vizualizaciji gozdov	5
Iztok Fister Jr. - Odkrivanje citatnih kartelov	6
Dušan Fister - Vizualizacija kolesarskega treninga	7
Primož Bencak - Izdelava MIDI klaviatur	8
Uroš Mlakar - Diferencialna evolucija in umetnost	9
Tilen Škrinjar - Detekcija jezika v šumnih besedilih s pomočjo hibridnih metod	10
Jani Dugonik - Sistemi za samodejno razpoznavo govora	11

# Analiza indikatorjev uspešnosti v večkriterijski optimizaciji

Miha Ravber  
*Univerza v Mariboru*

## **Povzetek**

Živimo v času, ko se že skoraj na vsakem koraku srečujemo s problemi, ki imajo več pomembnih kriterijev. Zato je za rešitev takšnih problemov potrebno sklepati kompromise. Izbira kompromisov pa je mnogokrat zahteven in dolgotrajen postopek. Na področju računalniške optimizacije pravimo, da imamo pri reševanju problemov z več kriteriji opravka z večkriterijsko optimizacijo. Za le-to je bilo razvitih mnogo različnih algoritmov. Vendar je primerjava rešitev algoritmov večkriterijske optimizacije veliko večji izziv kot pri algoritmih za enokriterijsko optimizacijo. Pri večkriterijski optimizaciji ne dobimo eno samo rešitev, temveč nabor rešitev, imenovan Pareto optimalna fronta. Zaradi težavnosti reševanja večkriterijskih problemov algoritmi običajno vrnejo aproksimacijo Pareto fronte. Te aproksimacije je potrebno oceniti preden jih lahko primerjamo. Najpogostejša metoda za ocenitev je uporaba indikatorjev uspešnosti. Indikator uspešnosti vrne oceno, ki odraža določen vidik kakovosti za podano aproksimacijo. Razvitih je bilo mnogo različnih indikatorjev uspešnosti, ki ocenjujejo različne vidike kakovosti. En sam indikator ne more zanesljivo oceniti uspešnost algoritma. Torej, kadar primerjamo algoritme večkriterijske optimizacije, je potrebno uporabiti različen nabor indikatorjev uspešnosti. Ker izbira indikatorja močno vpliva na oceno aproksimacije in posledično na uspešnost algoritma, je indikatorje potrebno predhodno analizirati. V tem delu bomo uporabili šahovski sistem rangiranja, prilagojen za primerjavo evolucijskih algoritmov večkriterijske optimizacije za analiziranje vpliva indikatorjev uspešnosti na rang algoritma. Primerjali bomo 11 različnih indikatorjev uspešnosti. Izvedli bomo 11 turnirjev in pri vsakem bomo uporabili različen indikator uspešnosti za ocenitev aproksimacijskih Pareto front. Kot rezultat vsakega turnirja bomo dobili lestvico vseh sodelujočih algoritmov. Za vsak algoritem bo podan njegov rang in odklon ranga, na podlagi katerega bomo izračunali interval zaupanja. Indikatorji uspešnosti bodo kategorizirani v skupine glede na to, kako podobno so rangirali algoritme. Pri podobnosti rangiranja bo upoštevan tudi interval zaupanja. Torej, če dva algoritma med seboj nimata signifikantne razlike, imata enak rang. Eksperiment bo izveden enkrat na desetih sintetičnih problemih iz tekmovanja CEC 2009 in enkrat na realnem problemu CITO. Tako bomo dobili nabor indikatorjev, ki podobno rangirajo na različnih problemih. Indikatorji, ki podobno rangirajo na različnih problemih, so bolj robustni in primernejši za uporabo.

# Prikaz senc pri vizualizaciji gozdov

Štefan Kohek  
*Univerza v Mariboru*

## **Povzetek**

Za obdelavo in analizo prostorskih podatkov pogosto potrebujemo vizualizacijo prostranih področij. V primeru prikaza simulacije rasti dreves potrebujemo predvsem 3D prikaz dreves in okolja, v katerem se drevesa nahajajo. Na vključitev prikaza senc, ki izboljša tako kakovost predstavitve kot tudi občutek globine, pa se pogosto pozabi. Najpogosteje uporabljen pristop za prikaz senc v igrah deluje na podlagi teksture senc, saj je enostaven za implementacijo, vendar ima mnoge težav, ki se opazijo še posebej pri prikazu večjih področij. Osnoven pristop namreč zahteva veliko teksturo senc, ki zajema celotno sceno. V primeru manjših tekstur postanejo sence, ki so bližje kameri, slabše kakovosti. Za izboljšanje kakovosti je tako potrebno povečati ločljivost, kar pa opazno poveča prostorske zahtevnosti. Novejši pristopi to težavo rešujejo na več načinov. Nekateri uporabljajo več tekstur senc, kjer za bližnja področja uporabljajo teksture z večjo ločljivostjo kot za oddaljena področja. Slabost tega pristopa je v tem, da zahteva večkratni zagon grafičnega cevovoda. Preostali pristopi delujejo samo z eno teksturo senc tako, da sliko popačijo na takšen način, da so bližnja področja predstavljena z več piksli kot oddaljena. Vendar pa to prinaša nove težave ali zahteva celo spremembo grafičnega cevovoda. Dodatne slabosti osnovnega pristopa so v tem, da v osnovi ne deluje na podlagi fizikalnih pravil. Oddaljeni objekti ponavadi tvorijo bolj mehke sence kot bližnji objekti, kar pa pristop s teksturami senc v osnovi ne more prikazati. V tem delu tako pregledamo različne pristope za izboljšanje prikaza senc s teksturami senc za vizualizacijo prostranih področij. Glede na to, da je večina pristopov za prikaz senc implementiranih znotraj obstoječih ogrodij za vizualizacijo, jih je potrebno ustrezno prilagoditi. Zato še pregledamo tehnologije, ki omogočajo enostavno vključitev obravnavanih pristopov v obstoječo vizualizacijo.

# Odkrivanje citatnih kartelov

Iztok Fister Jr.  
*Univerza v Mariboru*

## **Povzetek**

Dandanes postaja objavlanje znanstvenih člankov že skoraj tekmovanje. Raziskovalci, ki delujejo na različnih raziskovalnih inštitucijah, morajo za obstanek objavljati obilo izsledkov svojega raziskovalnega dela. Nemalo objav pa za obstanek ni več dovolj, ampak v veliki meri na to vplivajo tudi zloglasni citati, ki čedalje bolj prodirajo v veljavo. Številne citate pridobi avtor, ki objavlja prodorna ter za širšo javnost zanimiva dela. A kaj hitro zaključimo z mislijo, da to seveda v vsekakršni raziskovalni branži ni mogoče in da pogoji za vse raziskovalce niti malo ne sovpadajo. Zato se je, podobno kot v športu, tudi pri raziskovalcih pojavil ene vrste "akademski doping". Raziskovalci so se namreč začeli povezovati v ogromne združbe, t.i. "citatne kartele", kjer med sabo citirajo dela in s tem umetno povečujejo število citatov, kakor tudi neposredno vplivajo na uspešnost vsakega raziskovalca. V naši predstavitvi predstavimo algoritem za detekcijo citatnih kartelov, ki temelji na semantičnem spletu. Avtorje in članke predstavimo s pomočjo relacij, ki jih medsebojno povežemo in nato na podlagi teh s sklepalnikom poiščemo posamezne kartele. Začetni rezultati na sintetično generiranih podatkih so zelo spodbudni, zato bomo v predstavitvi predstavili še idejo, kako uporabiti algoritem na realnih podatkih iz različnih akademskih baz.

# Vizualizacija kolesarskega treninga

Dušan Fister

*Univerza v Mariboru*

## **Povzetek**

Analiza športnega treninga predstavlja odskočno desko za povečanje telesnih zmogljivosti športnika. Izvesti kakovostno analizo treninga je bilo nekdanj, zaradi pomanjkanja podatkov o športnih treningih, nemogoče opravilo, a je pojav, t.i. športnih sledilcev (angl. Sport-tracker), omogočil ogromno količino teh podatkov ter analizo približal širšemu krogu uporabnikov. Poseben pečat so sledilci pustili v kolesarstvu, saj so z razvojem različnih senzorjev omogočili beleženje slehernega parametra, ki se lahko pojavi pri treningu kolesarja in nadzor širokega spektra trenutnih in povprečnih kolesarskih parametrov. Povprečna in trenutna hitrost, kilometrina treninga, povprečni in trenutni srčni utrip, povprečna in trenutna moč in višinski metri so le nekaj izmed njih, saj se z vsako posodobitvijo kolesarske opreme ponujajo novi senzorji, ki omogočajo sledenje novih in novih parametrov športnega treninga. Sledilci beležijo podatke o kolesarskem treningu v datoteko, ki jo lahko kolesar po koncu treninga naloži na spletno stran njegovega proizvajalca. Spletna stran pridobljene podatke obdela in za vsak trening prikaže obilo numeričnih podatkov, številne grafe, tabele in koeficiente o treningu, ocenjevanja in prognoze zmogljivosti, čas spanja in dnevno težo ter skupne podatke skozi določeno časovno obdobje. Namen vseh prikazanih podatkov služi analizi treninga, kar pomeni, da lahko kolesar datoteko pregleda, razbere ter svoje podatke ovrednoti. Tako lahko neposredno spremlja svojo formo in vpliva na razvoj nadaljnih telesnih zmogljivosti. Glede na to, da je za ovrednotenje teh podatkov potrebno veliko izkustvenega znanja, se lahko kolesar začetnik v zmedu vseh števil hitro izgubi. To pomeni, da iz analize ne pridobi želenih informacij in posledično ne vpliva na razvoj svojih sposobnosti. Naš pristop k izboljšanju omenjene situacije in lažjemu razumevanju kolesarskega treninga je implementacija avtomatskega ovrednotenja podatkov in naknadne vizualizacije treninga. Vizualizacija treninga pomeni grafični prikaz ovrednotenih podatkov na človeku čim bolj razumljiv način. V skladu s tem je naš pristop razpet med dvema deloma, tj. prvim, ki samodejno analizira podatke o treningu, in drugim, ki ovrednoti in vizualizira najpomembnejše izmed njih. Analiza je izvedena z analizatorjem datotek, vrednotenje pa poteka z uporabo datoteke osebnih značilnosti (oz. osebno datoteko karakteristik), ki jo mora kolesar izpolniti pred prvo uporabo. Datoteka osebnih značilnosti priskrbi osnovne telesne podatke, npr. maksimalni srčni utrip, povprečno moč na eno uro, predvideno normo dnevne kilometrine ipd. in se z nadaljno uporabo avtomatsko posodablja. Rezultat vizualizacije je avtomatsko generirana slika ovrednotenih podatkov kolesarskega treninga, iz katere je moč razbrati intenziteto pridobljenih kolesarskih podatkov, vremenske razmere in po kasnejšem posredovanju uporabnika tudi počutje kolesarja. Praktično to pomeni, da generirana slika ponazarja podrobno analizo opravljenega treninga, zato je pristop vizualizacije uporaben ne samo za kolesarje začetnike, ampak tudi za profesionalne kolesarje.

# Izdelava MIDI klaviatur

Primož Bencak  
*Univerza v Mariboru*

## **Povzetek**

Pojav mikroročunalnika Raspberry Pi je na področju računalništva povzročil pravo malo revolucijo. Ob svoji majhni ceni in velikosti, v sebi skriva mogočen potencial, ki smo ga poskušali odkriti med izdelavo vse-v-enem MIDI klaviatur. Naša naloga je bila sestaviti MIDI klaviature, ki so samostojno delujoče (ne potrebujejo osebnega računalnika) in omogočajo osnovno manipulacijo z MIDI datotekami. Za sprejemanje ukazov uporabnika, kot so detekcija not in spreminjanje višine le-teh, smo uporabili mikrokrmilnik MSP430F5529 Launchpad podjetja Texas Instruments, za sintetizacijo zvoka pa omenjeni mikroročunalnik Raspberry Pi 2. Da bi uporabniku omogočili pregledno in enostavno upravljanje s klaviaturami, smo na čelni plošči dodali dvovrstični LCD zaslon in zaslon, občutljiv na dotik, ki je povezan z mikroročunalnikom Raspberry Pi in LED indikatorje, ki se aktivirajo ob pritisku na gube. Klaviature so namenjene igranju v prostem času in kot zabaven pripomoček k učenju. Napaja jih polnilna baterija litij-polimer, ki omogoča avtonomnost do približno 8 ur. V lesenem ohišju klaviatur je vgrajen tudi ojačevalec in dva zvočnika, kar odstrani potrebo po dodatni opre. Pritisk gumba na tipkovnici sproži generiranje MIDI note, kar opravi mikrokrmilnik MSP430F5529. Ta prav tako spreminja višino not (transponiranje) in skrbi za prikaz spremembe nastavitve na dvovrstičnem LCD zaslonu. Ta je z mikroročunalnikom Raspberry Pi povezan z USB kablom, preko navideznih serijskih vrat. Program ttyMIDI, ki je nameščen na mikroročunalniku, skrbi za preverbo in posredovanje poslanih MIDI not z mikrokrmilnika naprej v uporabniški program za sintetizacijo zvoka (npr. Hydrogen, Swami). Od tukaj naprej je uporabniku prepuščena izbira programa in zvoka, ki bi ga rad slišal ob igranju na klaviature. Na mikroročunalniku nam je poleg igranja klaviatur omogočeno še vse ostalo, kar ponuja omenjeni operacijski sistem (brskanje po spletu, ogled videov, ipd.) Klaviature imajo še dodaten USB priključek, v katerega lahko priklopimo USB medij, ali pa iz njega napajamo kompatibilne naprave (zvočniki). Stikalo na zadnji strani omogoča preklap med vgrajenim ojačevalcem in dodatnimi zvočniki ali line-out izhodom. Vsa elektronska vezja in ohišje so bila ročno izdelana in narisana v programu Eagle in Target3001! ter Solidworks. Namen projekta je bil prikaz uporabe mikroročunalnika v vsakodnevnih aplikacijah.



# Diferencialna evolucija in umetnost

Uroš Mlakar  
*Univerza v Mariboru*

## **Povzetek**

Algoritmi računske inteligence pridobivajo na vedno večji popularnosti, saj praktično prodirajo že na vsa aplikativna področja računalništva. Pojavlja se tudi trend pojavljanja teh algoritmov v ostalih disciplinah kot sta šport in umetnost. Naše delo se osredotoča predvsem na umetnost, kjer se v raziskovalni srenji pojavlja izraz "evolutionary art". Naredili smo kratek pregled področja in skušali zbrati vse algoritme, ki se pojavljajo v omenjeni disciplini. Hkrati predstavimo tudi našo metodo, ki temelji na algoritmu diferencialne evolucije. Cilj predlaganega algoritma je predstaviti sliko kot množico trikotnikov, ki bodo kot celota najboljše predstavili vhodno sliko. Testiranje smo izvedli nad 7 slikami različnih težavnosti, kjer se naša metoda na nekaterih odraža bolje kot na ostalih. Rezultati kažejo, da je diferencialna evolucija primeren algoritem za uporabo v umetnosti, hkrati pa v zaključku navedemo še vse posredne aplikacije omenjenega dela.

# Detekcija jezika v šumnih besedilih s pomočjo hibridnih metod

Tilen Škrinjar  
*Univerza v Mariboru*

## **Povzetek**

Dandanes na svetovnem spletu zasledimo kratka besedila, ki so mnogokrat slovnično nepravilna. Takšna besedila najdemo na forumih, v komentarjih in drugje. Pri tem se pojavijo težave pri avtomatski detekciji jezika. Še težje pa razpoznamo jezike, kateri so si med seboj podobni oziroma spadajo v isto jezikovno skupino. Osredotočili smo se na skupino slovenskih jezikov. Za razpoznavo jezikov je bilo razvitih že več algoritmov, katere smo podrobneje pregledali in preučili. Za učni korpus smo uporabili dvajset najpogostejših besed posameznega jezika in nekaj značilnih črk za posamezni jezik. Uporabili smo devet jezikov. Pri testnem korpusu smo za posamezni jezik izbrali petindvajset besedil najdenih na forumih na svetovnem spletu. Testno besedilo klasificiramo glede na najboljšo kombinacijo znakov in besed v jeziku. Uspeh razpoznave smo analizirali s spreminjanjem različnih parametrov. Pri tem smo ugotavljali različne soodvisnosti med posamezni parametri. V to analizo smo vključili tudi druge konstrukte, kot je lematizator. Na koncu smo načrtovali lasten algoritem in analizirali posamezne parametre. Pri tem smo uporabili znake, besede in pa utežen vrstni red besed pri posameznem jeziku. Rezultate smo izboljšali za 0.4 %.

# Sistemi za samodejno razpoznavo govora

Jani Dugonik  
*Univerza v Mariboru*

## **Povzetek**

Govor, kot način komunikacije, predstavlja kompleksen vzorec, ki je podvržen številnim dejavnikom, kot npr. fiziološke karakteristike govorca, okolje v katerem poteka govor, čustveno razpoloženje govorca,.. in je zato močno variabilen. Razumevanje govora pri človeku predstavlja bistveno komponento v celotnem postopku razpoznavanja. Zaradi tega tudi človek izredno nezanesljivo razpoznavlja glasove nekega jezika, ki ga ne razume in ravno tako slabše razpoznavlja v lastnem jeziku nesmiselne besede (besede, ki ustrezajo akustičnim zakonitostim jezika, ne nosijo pa nobenega pomena). V ta namen so začeli iskati cenejše in hitrejša načina za razpoznavo govora – sistem za samodejno razpoznavo govora. Ti sistemi se uporabljajo v avtomobilskih sistemih, medicini, vojski, telefoniji, robotiki, itd. Dva pomembnejša razloga zakaj se je toliko raziskovanja posvetilo v problem razpoznavanja govora sta dostopnost za gluhe in naglušne, in znižanje stroškov z avtomatizacijo. Slabosti teh sistemov pa so možnost nezanesljivega delovanja, potrebno je uvajanje uporabnikov na nov način komunikacije, zmanjšanja števila delovnih mest, in pa te sisteme (razen najpreprostejših) je potrebno vsaj v določeni meri razviti za vsak jezik posebej. Samodejni sistem, ki bi razpoznaval in razumel govor brez dodatnih omejitev, ki zadevajo okolje, področje komunikacije, število možnih besed, je še vedno nedosežen cilj.