

# Lietuvos informatikos olimpiados darbų vertinimas programinės įrangos kokybės modelio požiūriu

## Jūratė Skūpienė

Matematikos ir informatikos instituto jaunesnioji mokslo darbuotoja  
 Institute of Mathematics and Informatics, Younger Research Fellow  
 Akademijos g. 4, LT-08663 Vilnius  
 Tel. (8 5) 210 97 32  
 El. paštas: jurate.sk@gmail.com, <http://ims.mii.lt/ims/asmen/sjurat.html>

*Mokinių informatikos olimpiados – tai programavimo konkursai, per kuriuos metų dalyviai turi sukurti tinkamą ir efektyvų algoritmą, jį realizuoti kuria nors leidžiama programavimo kalba. Sprendimai olimpiadose vertinami automatiškai testuojant programas, taip pat vertinamas algoritmo aprašymas ir programavimo stilius. Pati vertinimo schema jau daugelį metų yra susiklosčiusi remiantis tradicijomis, metodiniais samprotavimais, tačiau nėra mokliškai pagrįsta. Šis straipsnis – tai pirmasis žingsnis, siekiant mokliškai pagrįsti vertinimą informatikos olimpiadose. Straipsnyje į dalyvio sukurtą programą žvelgiama programinės įrangos kokybės reikalavimų standarto ISO-9126-1 požiūriu ir analizuojama, kiek dabartinė vertinimo schema atitinka šį kokybės modelį.*

## Įvadas

Lietuvos mokinių informatikos olimpiados (LitIO) yra individualios programavimo (algoritmavimo) varžybos vidurinių mokyklų mokiniams (Lietuvos ..., 2008; Dagienė, 2007). LitIO organizuojama remiantis Pasaulinių informatikos olimpiadų (International Olympiads in Informatics – IOI) modeliu (International..., 2008a; International..., 2008b), kuriuo remiasi daugelis nacionalinių bei regioninių konkursų (Diks, 2007; Cerchez, 2008; Dagienė, 2004). Tad straipsnis aktualus platesniame kontekste, nes tos pačios problemos, pirmiausia vertinimo, būdingos ir kitoms olimpiadoms.

LitIO pateikiamos algoritmavimo užduotys. Norėdami jas išspręsti, dalyviai turi pasitelkti algoritmų ir duomenų struktūrų žinias, programavimo ir derinimo įgūdžius, kūrybiškai visa tai taikydami sukurti uždavinį sprendžiantį algoritmą ir jį realizuoti leidžiama programavimo kalba (šiuo metu Paskaliu, C/C++). Toliau straipsnyje terminas *sprendimas* bus suprantamas kaip uždavinį

sprendžianti programa, kurioje komentarų pavidalu įrašytas sprendimo aprašymas. Sprendimai pateikiami vertinimui naudojant *Programavimo varžybų aptarnavimo sistemą* (PVAS) (Mareš, 2007). Ši sistema numato saityno sąsają sprendimams pateikti ir įvairias olimpiadai valdyti reikalingas paslaugas.

Vertinimas informatikos olimpiadose yra vienas opiausių su olimpiada susijusių klausimų, pastaruoju metu analizuojamų įvairiuose straipsniuose (Vasiga et al., 2008; Verhoeff, 2006; Forišek, 2006). Vertinimo schema, naudojama Lietuvos informatikos olimpiadose paremta tradicijomis (jau įvyko 20 olimpiadų), metodiniais samprotavimais ir IOI stebimomis tendencijomis (Dagienė, 2007). Tačiau iki šiol nebuvo bandymų ją mokliškai pagrįsti.

## Sprendimo (programos) kokybės samprata

Kalbant apie vertinimą olimpiadose, kyla diskusijų, kuri sprendimo dalis yra svarbesnė: ar pats algoritmas (*kompiuterių mokslo sritis*), ar

gebėjimas jį realizuoti programavimo kalba (*informatikos inžinerijos sritis*). Svarbiausias informatikos olimpiados tikslas – pasivaržyti sprendžiant problemas\*, t. y. kūrybiškai ir originaliai taikant turimą patirtį ir žinias naujomis aplinkybėmis. Šis tikslas lemia tai, kad gebėjimas sukurti teisingą ir efektyvų algoritmą yra svarbesnis nei gebėjimas jį realizuoti, tačiau informatikos olimpiadose abu šie dalykai yra glaudžiai susiję ir šiuo metu galutinis vertinimui pateikiamas darbas (programa) yra artimesnis programinės įrangos (PI) nei teorinei algoritmo sampratai. Šiame straipsnyje į sprendimą ir jo vertinimą pažvelgsime būtent tokiu požiūriu.

Prieš aptariant įvairias charakteristikas, nusakančias PI kokybę, būtina atsakyti į klausimą, kas tai yra kokybė. R. W. Hoyer (2001) nusako dvi požiūrių kryptis, kuriomis bandoma apibrėžti kokybę. Tai *atitiktis specifikacijai* (produktas laikomas kokybišku, jeigu jo išmatuojamos charakteristikos atitinka iš anksto numatytus reikalavimus šiam produktui) arba *atitiktis klientų poreikiams* (šiuo atveju kokybė yra gebėjimas atitikti kliento poreikius ir ji nėra siejama su matuojamomis charakteristikomis).

P. Berander (2005) apžvelgia įvairias kokybės sampratas. P. B. Crosby požiūriu, kokybė nėra prabanga ar elegancija, o griežta atitiktis reikalavimams. Kokybės standartas turi būti labai aiškus, ir požiūris „beveik gerai“ nėra priimtinas. Pasak W. E. Deming kokybė yra atitiktis vartotojo poreikiams, tačiau sunkumas nusakant kokybę atsiranda, kai reikia šią atitiktį užrašyti išmatuojamomis charakteristikomis. Jis teigia, kad kokybė nusakoma tik *agento*, t. y. konkretaus kokybės vertintojo, požiūriu. A. V. Feigenbaumas teigia, kad kokybę nustato tik vartotojas naudodamas produktą, tačiau jo poreikiai yra kintantys, tad ir kokybės sąvoka kintanti. Ishikawa taip pat yra tos nuomonės, kad kokybė yra atitiktis vartotojo poreikiams, kartu pabrėžia, kad tarptautiniai standartai (ISO, IEEE) turi trūkumų ir nepakankamai greitai reaguoja į kintančius vartotojų poreikius. W. A. Shewart į kokybę žiūri dviem požiūriais,

t. y. kaip į objektyvią realybę, nepriklausomą nuo vartotojų egzistavimo, ir kaip į subjektyvią, perspektyvą priklausomą nuo individų minčių, pojūčių, atsiradusių dėl objektyvios realybės.

Apibendrinant reiktų pasakyti, kad kokybė vis dėlto yra ne absoliuti, o konstruojama ir kintama sąvoka, todėl negalima kalbėti ir apie absoliučią olimpiadinio sprendimo (programos) kokybę. Du svarbiausi požiūriai (atitiktis specifikacijai ir klientų poreikiams) susipina kalbant apie olimpiados sprendimų vertinimą, nes ir specifikacijos sudarytojai (sprendimo reikalavimai), ir klientai yra ta pati žmonių grupė (t. y. vertinimo komisija).

Buvo sukurta įvairių PI kokybės modelių. Vienas žinomiausių pirmųjų modelių buvo McCall (1977) kokybės modelis, kuris numatė tris PI kokybės nusakymo kryptis: galimybė keisti PI, pritaikyti ją prie naujų aplinkų ir PI vykdomosios charakteristikos. Buvo sukurtas hierarchinis medis, kuris numatė faktorius (t. y. išorinis požiūris į PI – ką mato vartotojas), kokybės kriterijus (vidinis požiūris – ką mato PI gamintojas) bei metrikas, skirtas įvairiems kriterijų aspektams matuoti. Šiuo metu labiausiai paplitęs standartas yra ISO-9126 PI vertinimo modelis (ISO 2001), kuris buvo sukurtas remiantis ankstesniais, iš jų ir McCall darbais.

Požiūrį į vertinimui pateiktą sprendimą iš esmės lemia ekspertų pasirinkta kokybės samprata ir konkretus kokybės modelis. Dažnai pasitaiko situacija, kai kyla diskusijų dėl vertinimo kokybės kriterijų ir konkrečių metrikų nesusitarus, o kaip suprantama kokybė ir koku kokybės modeliu remiamasi. Analizuodami sprendimų vertinimą informatikos olimpiadose remsimės tarptautiniu standartu ISO-9126 PI vertinimo modeliu. Viena priežasčių, dėl ko pasirinktas šis modelis analizuojant olimpiadų vertinimo schemą, yra ta, kad egzistuoja didelė olimpiadinės bendruomenės požiūriu įvairovė, tad visuotinai pripažintas standartas suteikia svorį pagrindžiant vertinimo schemą. Pasirėmus šiuo standartu vėliau kokybės modelį galima modifikuoti, t. y. pritaikyti prie labai konkrečių olimpiados poreikių.

\* Angl. *Problem Solving*

## **Vertinimo informatikos olimpiadose schema atsižvelgiant į PĮ ISO-9126-1 kokybės standartą**

Į olimpiados dalyvio sukurtą programą pasižiūrėsime PĮ keliamų kokybės reikalavimų požiūriu. Šiuos reikalavimus nusako ISO-9126 standartas, kurį sudaro keturios dalys: kokybės modelis, išorinės metrikos, vidinės metrikos ir naudojimo kokybės metrikos (ISO-9126, 2009). Kokybės modelis yra pirmoji standarto dalis (ISO-9126-1) ir naudojamas vertinti programai, kai pateiktas jos pradinis tekstas. Kitos trys dalys svarbesnės PĮ kokybei užtikrinti ir PĮ kūrimo procesui tobulinti (ISO-9126, 2009; Software, 2007).

ISO-9126-1 PĮ kokybės modelis numato šešias PĮ charakteristikas: funkcionalumą, patikimumą, panaudojamumą, efektyvumą, prižiūrimumą, perkeliamumą. Lentelėje apžvelgsime, kaip atrodo dabartinė vertinimo schema pagal šias charakteristikas.

Vertinant olimpiadų dalyvių sprendimus taškai skiriami už tris charakteristikas iš šešių: funkcionalumą, efektyvumą ir pataisomumą. Olimpiados rengėjų nuostata, kad dalyviai turi sutelkti dėmesį į algoritmą ir jo realizaciją, ku mažiau dėmesio skirti techniniams aspektams, lemia, kad patikimumas, kuris yra svarbus, kalbant apie nuolatiniam vartojimui skirtą PĮ, nėra vertinamas. Panaudojamumas nėra tiesiogiai vertinamas, tačiau tai yra būtina kartelė, kurią reikia įveikti, kad sprendimas būtų priimtas vertinti. Perkeliamumas turi daug įvairių aspektų ir dalis jų techniškai yra susiję su kompiliatorių versijomis, naudojamomis bibliotekomis ir pan. Olimpiadose visuomet būdavo siekiama, kad techninių subtilybių išmanymas nebūtų lemiamas dalykas, tad visai suprantama, kad ši charakteristika nevertinama.

### **Kokybės modelio charakteristikų – funkcionalumo, efektyvumo ir pataisomumo – vertinimas**

Svarbiausia kokybės modelio charakteristika yra funkcionalumas (informatikos olimpiadų kontekste vadinamas *teisingumu*), ir svarbiau-

sias olimpiadose nuolat keliamas klausimas – kaip įvertinti sprendimo teisingumą (Verhoeff, 2006; Forišek, 2006).

Šiuo metu olimpiadose teisingumą priimta vertinti automatiškai testuojant programas. Iš anksto parengiami įvairaus sudėtingumo testai, nagrinėjantys skirtingus atvejus. Programa vykdoma su kiekvienu testu ir už testą skiriamai balai, jei programa sėkmingai užbaigia darbą neviršydama laiko ir atminties ribojimų ir pateikia teisingą rezultatą. Teisinga laikoma ta programa, kuri įveikia visus testus, tačiau dalį testų įveikusi programa surenka dalį už teisingumą skiriamų taškų, kitaip sakant, taikomas dalinis teisingumo vertinimas. Kadangi dalyviai neprivalo pateikti realizuoto algoritmo teisingumo įrodymų (vidurinių mokyklų mokiniais pateikti įrodymus mokslinė kalba būtų per daug sudėtinga) ir nėra privalomų programavimo stiliaus reikalavimų, automatinis testavimas laikomas vienintele efektyvia ir priimtina alternatyva informatikos olimpiadose (Cormack, 2006a).

Funkcionalumą kokybės modelio požiūriu įprasta vertinti logine funkcija, t. y. arba jis yra, arba ne (ISO-9126, 2009). Tokia vertinimo sistema taikoma komandinėse studentų informatikos olimpiadose ACM ISPC (Cormack, 2006), kur už ne iki galo išspręstą uždavinį taškai neskiriami. Tiesa, ne taip kaip mokinių olimpiadose: pateikiami ne du ar trys, o dešimt vienuolika uždavinių. Kalbant apie dalinį teisingumo vertinimą susikerta problemų sprendimo, edukologiniai ir inžineriniai požiūriai. PĮ kokybės požiūriu, jei programa vienais atvejais pateikia teisingą rezultatą, o kitais – klaidingą, ji nėra funkcionali. Tad dalinio vertinimo neturėtų būti. Edukologiniu požiūriu loginis vertinimas, kai dalyviai sprendžia tik du ar tris uždavinius, būtų netinkamas, nes mažintų dalyvių motyvaciją (Leeuwen, 2005; Skūpienė, 2006). Buvo analizuoti trijų olimpiadų konkrečių uždavinių sprendimai ir nustatyta, kad vertinant loginiu principu ne mažiau kaip pusė dalyvių už tuos uždavinius būtų visai negavę taškų.

Problemų sprendimo požiūriu toks dalinis vertinimas, koks jis taikomas šiuo metu, kelia daug klausimų ir abejonių. Pasaulinėse olimpiadose dalinis teisingumo vertinimas realizuojamas

*Lentelė. Vertinimo schemas suskaidymas pagal ISO-9126-1 kokybės modelio charakteristikas*

<i>Charakteristika</i>	<i>Apibūdinimas</i>	<i>Ar vertinama informatikos olimpiadose</i>	<i>Pastabos</i>
<i>Funkcionalumas</i>	Skirta įvertinti, kaip PĮ atlieka savo svarbiausių paskirtį, t. y. ar realizuotos visos reikalaujamos funkcijos	Nuo 40 % iki 100 % taškų*	Svarbiausia vertinamos programos paskirtis – išspręsti uždavinį, t. y. suskaičiuoti ir pateikti reikalaujamą rezultatą; olimpiadoje šią charakteristiką įprasta vadinti <i>teisingumu</i> ; ji tikrinama automatiškai testuojant sprendimus; tačiau keliamas klausimas, ar testavimas pakankamai gerai patikrina šią charakteristiką.
<i>Patikimumas</i>	Skirta įvertinti jau paleistos sistemos gebėjimui patikimai veikti	Statistika renkama, tačiau taškai neskiriami	Uždavinių sąlygose aiškiai nurodomas duomenų formatas ir rekomenduojama netikrinti jokių nenumatytų atvejų; atlikusi skaičiavimus programa turi pati užbaigti darbą; olimpiados vykdymo tvarkoje neregamentuota, kaip turi elgtis programa, jei kurioje nors situacijoje ji neranda sprendinio; dažnai tokiose situacijose ji stringa, lūžta arba pateikia neteisingą rezultatą. Nors iš rezultatų matoma, kiek kartų programa nebaigė darbo sėkmingai, tačiau tai nėra atskirai matuojama ir vertinama, o prilyginama klaidingai suskaičiuotam rezultatui; ši charakteristika nėra akcentuojama olimpiadose ir nėra vertinama sprendimuose, nors minimali statistika apie nesėkmes (lūžimą ar užstrigimą) surenkama.
<i>Panaudojamumas</i>	Skirta įvertinti, ar vartotojui lengva (išmokti) naudotis PĮ	Atskirai taškai neskiriami; tačiau vertinti siunčiamos tik reikalavimus atitinkančios programos	Visi išorinės sąsajos reikalavimai yra griežtai reglamentuoti, o patikrinti, ar jų laikomasi, padeda PVAS. Ši charakteristika taip pat nėra matuojama vertinant dalyvių programas, tačiau reikalavimų netenkinančios programos nevertinamos.
<i>Efektyvumas</i>	Skirta įvertinti PĮ efektyvumui laiko ir kitų išteklių atžvilgiu	Nuo 30 % iki 60 % taškų;	Sąlygoje numatomi atminties ir vykdymo laiko ribojimai, kurių negali viršyti programa; šių išteklių panaudojimas nėra vertinamas tiesiogiai (t. y. per trumpesnį laiką baigusi darbą ar mažiau atminties sunaudojusi programa nesurenka daugiau taškų); tačiau dalis testų skiriami efektyvumui vertinti, t. y. duomenų sudėtingumas juose palaipsniui didinamas ir taip sprendimai suskirstomi į kategorijas pagal sprendimo efektyvumą laiko ir atminties atžvilgiu.
<i>Pataisomumas</i>	Skirta įvertinti, ar lengva modifikuoti PĮ	Iki 10 % taškų	Ši charakteristika vertinama skiriant balus už <i>programavimo stilių</i> ; yra parengtos programavimo stiliaus gairės, kurios yra bendros ir nereikalauja prisirišti prie konkretaus stiliaus, tačiau atkreipia dėmesį į bendrus reikalavimus; stilius vertinamas naudojant Likerto skalę, o ne sumuojant taškus už konkrečius stiliaus charakteristikas; siekiant išvengti atvejų, kad nebūtų skiriami taškai už uždavinio nesprenžiančios programos stilių, vertinamas tik programų, už testus surinkusių ne mažiau kaip 50 % taškų, stilius.
<i>Perkeliamumas</i>	Skirta įvertinti, ar lengva perkelti PĮ iš vienos platformos į kitą	Nevertinama	Atskirai ši charakteristika nėra vertinama; tačiau netiesiogiai ir labai minimaliai vertinama skiriant balus už programavimo stilių – per didelį prisirišimą prie konkrečios aplinkos vertinamas neigiamai**.

\* Konkretaus uždavinio vertinimo schema sudaroma prieš olimpiadą; tačiau skirtingiems uždaviniams taškai, skiriami už vieną ar kitą aspektą, gali skirtis.

\*\* Lietuvos olimpiadose susiklosčiusi savotiška situacija, kai dalis dalyvių dirba *Windows* operacinėje sistemoje, o PVAS veikia ir sprendimus testuoja *Linux* operacinėje sistemoje; tad per didelis prisirišimas prie konkrečios operacinės sistemos gali kainuoti dalyviams laiko ir pastangų, kol jie pasiekia, kad programa veiktų ir jų kompiuteryje, ir vertinimo mašinoje.

suskaidant uždavinį į mažesnius dalinius uždavinius ir atskirai tikrinant, ar išspręstas kiekvienas dalinis uždavinys. LitIO dalinis vertinimas realizuojamas skiriant taškus už kiekvieną įveiktą testą, tačiau užduotis nėra skaidoma į dalinius uždavinius. Svarbiausia priežastis ta, kad rašant sprendimą (algoritmą, programą) neefektyvu kurti keletą atskirų dalinių sprendimų (procedūrų), kurias sujungus būtų gaunamas viso uždavinio sprendimas. Dažniausiai konstruojamas vienas išsistinis sprendimas, apimantis visus arba tik kai kuriuos dalinius uždavinius. Viena klaida tokiaame sprendime, tiesiogiai nesusijusi su konkretaus dalinio uždavinio specifika, gali lemti, kad programa klaidingai skaičiuos kelius (ar kartais net visų) dalinių uždavinių sprendinius. Antroji problema, su kuria susiduriama vertinant teisingumą, yra pats vertinimo būdas – automatinis testavimas. Viena vertus, testavimas, kuris yra priimtinas kaip klaidų ieškojimo būdas, nėra laikomas patikimu teisingumo įrodymo metodu (Dijkstra, 1970). Kita vertus, jei programa neišveikia konkreto testo, bendru atveju nieko negalima pasakyti apie joje esančios klaidos mastą: tai gali būti arba smulki korektūros klaida, arba realizacijos klaida, arba didelė klaida pačiame algoritme. Toks vertinimas neužtikrina, kad algoritmiškai (problemų sprendimo požiūriu) lygiaverčiai sprendimai būtų vertinami panašiu taškų skaičiumi. Vienas siūlomų sprendimų – naudojant PVAS galimybes dar olimpiados metu pateikti kuo daugiau grįžtamojo ryšio (Vasiga, 2008).

Kaip minėta, funkcionalumas yra svarbiausia charakteristika ir kitų charakteristikų matavimas turi prasmę tik tuomet, kai programa yra funkcionali. Tad būtų logiška, jei už efektyvumą ir pataisomumą taškai būtų skiriami tik tuomet, kai programa surenka daug taškų už teisingumą. Pataisomumas (programavimo stilius) iš tiesų yra vertinamas tik jei programa surenka tam tikrą taškų skaičių ( $\geq 50\%$ ) už teisingumą, o efektyvumas vertinamas visais atvejais. Reikėtų atkreipti dėmesį, kad euristiniai (t. y. neteisingi) sprendimai dažnai esti efektyvesni nei teisingi sprendimai, tad vertėtų pagalvoti apie efektyvumo vertinimo susiejimą su teisingumo įver-

čiais. Viena priežasčių, kodėl tai nepadaryta iki šiol, – neišspręstas pirmiau minėtas trijų požiūrių į funkcionalumo vertinimą konfliktas. Mat dažnai parašyti teisingą ir neefektyvų sprendimą paprasta, o teisingą ir efektyvų – gana sudėtinga, tam reikia gerų problemų sprendimo ir programavimo įgūdžių.

## Išvados

Vertinant olimpiadinio uždavinio sprendimą reikėtų atskirti du aspektus – tai sukurtas algoritmas ir jo realizacija programavimo kalba (programa). Olimpiadinio uždavinio sprendimo (programos) kokybės samprata negali būti vienareikšmė – tai konstruojama sąvoka, priklausanti nuo konkreto pasirinkto kokybės modelio. Dabartinė vertinimo schema atitinka ISO-9126-1 standarto kokybės modelį, joje numatoma, kad vertinamos trys iš šešių šio kokybės modelio charakteristikų: funkcionalumas, efektyvumas ir pataisomumas. Patikimumas ir perkeliamumas nėra vertinami, panaudojimumui keliami konkretūs nesudėtingi reikalavimai, kuriuos būtina įvykdyti, kad programa būtų apskritai vertinama.

Pagal PĮ kokybės modelį svarbiausia charakteristika yra funkcionalumas ir jį vertinti yra keblu, nes susikerta edukologiniai, PĮ kokybės modelio ir problemų sprendimo (t. y. svarbiausio olimpiados tikslo) požiūriai į vertinimą. PĮ kokybės modelio požiūriu vertinimas turėtų būti loginis, siekiant išlaikyti dalyvių motyvaciją, jis turėtų būti dalinis, o problemų sprendimo požiūriu dabar egzistuojantis automatinis funkcionalumo matavimas apskritai nėra geras, nes visiškai nelygiaverčiai algoritmai gali būti įvertinti panašiu taškų skaičiumi (t. y. automatinis testavimas atskleidžia klaidos egzistavimą, tačiau neatskleidžia klaidos masto).

Apibendrinant galima pasakyti, kad iš esmės informatikos olimpiados sprendimų (programų) vertinimas pagal ISO-9126-1 kokybės modelį yra pagrįstas, tačiau matuojant (vertinant) funkcionalumą kyla prieštaravimas tarp skirtingų požiūrių ir tam būtina ieškoti kitų sprendimų ar kito pagrindimo.

## LITERATŪRA

- BERANDER, P. et al. (2005). Software Quality Attributes and Trade-offs. *Blekinge Institute of Technology: elektroninė knyga* [žiūrėta 2009 m. gegužės 20 d.]. Prieiga per internetą: <http://www.bth.se/tek/besq.nsf/pages/017bd879b7c9165dc12570680047aae2!OpenDocument>.
- CERCHEZ, E.; ANDREICA, M. I. (2008). Romanian national olympiads in informatics and training. *Olympiads in Informatics: Tasks and Training*, vol. 2, p. 37–47.
- CORMACK, G. (2006a). Random factors in IOI 2005 test case scoring. *Informatics in Education*, vol. 5, no. 1, p. 5–14.
- CORMACK, G.; KEMKES, G.; MUNRO, I.; VASIGA, T. (2006b) Structure, scoring and purpose of computing competitions. *Informatics in Education*, vol. 5, no. 1, p. 15–36.
- DAGIENĖ, Valentina; SKŪPIENĖ, Jūratė (2007). Contests in programming: quarter century of Lithuanian experience. *Olympiads in Informatics: Country Experiences and Developments*, vol. 1, p. 37–49.
- DAGIENĖ, Valentina; SKŪPIENĖ, Jūratė (2004). Learning by competitions: Olympiads in Informatics as a tool for training high grade skills in programming. *Iš Proceedings of 2nd International Conference Information Technology: Research and Education*. T. Boyle, P. Oriogun, A. Pakstas (Eds.). London, p. 79–83.
- DIJKSTRA, E. W. (1970). *On the reliability of programs*. Prieiga per internetą: <http://www.cs.utexas.edu/users/EWD/transcriptions/EWD03xx/EWD303.html>.
- DIKS, Krsztof; KUBICA, Marcin; STENCEL, Krzysztof (2007). Polish Olympiads in Informatics – 14 Years of Experience. *Olympiads in Informatics: Country Experiences and Developments*, vol. 1, p. 50–56.
- FORIŠEK, Michal (2006). On the suitability of tasks for automated evaluation. *Informatics in Education*, vol. 5, no. 1, p. 63–76.
- HOYER, R.W; HOYER, B. B. Y. (2001) What is Quality? *Quality Progress*, no 7, p. 52–62.
- International Olympiads in Informatics* (2008a). [žiūrėta 2009 m. gegužės 14 d.]. Prieiga per internetą: <http://www.ioinformatics.org/index.shtml>.
- International Olympiads in Informatics – Regulations* (2008b) [žiūrėta 2009 m. gegužės 14 d.]. Prieiga per internetą: <http://www.ioinformatics.org/rules/reg08.pdf>.
- ISO 9126 Software Quality Characteristics*. (2009) [žiūrėta 2009 m. gegužės 17 d.]. Prieiga per internetą: <http://www.sqa.net/iso9126.html>.
- ISO, International Organization for Standardization, ISO 9126-1:2001, Software engineering – Product Quality, Part 1: Quality model, 2001.
- KEMKES, G.; VASIGA, T.; CORMAK, G. (2006) Objective Scoring for Computing Competition Tasks. *Iš Informatics Education – the Bridge between Using and Understanding Computers: Proceedings of International Conference in Informatics in Secondary Schools – Evolution and Perspectives*, November 7–11, Lecture Notes in Computer Science. Springer-Verlag, p. 230–241.
- Van LEEUWEN, W.T. (2005). *A Critical Analysis of the IOI Grading Process with an Application of Algorithm Taxonomies*. Master's Thesis, Technische Universiteit Eindhoven, Faculty of Mathematics and Computing Science [žiūrėta 2009 m. gegužės 17 d.]. Prieiga per internetą: <http://alexandria.tue.nl/extra1/afstversl/wsk-i/leeuwen2005.pdf>.
- Lietuvos mokinių informatikos olimpiados sąlygos* (2008). Patvirtintos LMITKC direktoriaus 2008. XI.19. [žiūrėta 2009 m. gegužės 14 d.]. Prieiga per internetą: <http://www.lmitkc.lt/lt/informatikos>.
- MAREŠ, Martin (2007). Perspectives on Grading Systems. *Olympiads in Informatics: Country Experiences and Developments*, vol. 1, p. 124–130.
- MCCALL, J.A; RICHARDS, P.K.; WALTERS G.F. (1977). *Factors in Software Quality*. Vols I–III, Rome Air Development Center, Italy.
- SKŪPIENĖ, Jūratė (2009). Credibility Of Automated Assessment In Lithuanian Informatics Olympiads: One Task Analysis. *Pedagogika*, ISSN 1392-0340, priimta spausdinti 2009 metų numerįje.
- SKŪPIENĖ, Jūratė (2006). Programming Style – Part of Grading Scheme in Informatics Olympiads: Lithuanian Experience. *Iš Informatics Education – the Bridge between Using and Understanding Computers: Proceedings of International Conference in Informatics in Secondary Schools*. Information Technologies at School, November 7–11. Vilnius: TEV, p 545–552.
- Software Quality ISO Standards (2007). *Arisa – Controlling Software Quality* [žiūrėta 2009 m. gegužės 17 d.]. Prieiga per internetą: <http://www.arisa.se/compendium/node6.html>.
- VASIGA, T; CORMACK, G.; KEMKES, G. (2008). What do olympiads tasks measure? *Olympiads in Informatics: Tasks and Training*, vol. 2, p. 181–191.
- VERHOEFF, Tom (2006). The IOI is (not) a science olympiad. *Informatics in Education*, vol. 5, no. 1, p. 147–158.

## ASSESSMENT OF SOLUTIONS OF LITHUANIAN INFORMATICS OLYMPIADS FROM THE POINT OF VIEW OF SOFTWARE QUALITY MODEL

Jūratė Skūpienė

### Summary

Lithuanian Informatics Olympiads are algorithmic and programming competitions where the contestants are given algorithmic tasks and have to design and implement correct and efficient algorithms in one of the allowed programming languages (currently Pascal, C/C++). The current grading scheme consists of automated testing and manual assessment of verbal algorithm description and programming style. Two aspects of submission should be taken into account – algorithm (computer science area) and its implementation (computer engineering field). The final submission (program) delivered for grading is more likely to resemble piece of software rather than constructed algorithm. The paper focuses on analysing the grading of submissions from the point of view of Software quality model. Understanding of quality of a submission can't be homogenous as quality is a constructed notion dependant on a particular software quality model chosen by the experts (scientific committee of the Olympiad). The grading schema (except for verbal description of an algorithm) correspond ISO-9126-1 software quality model. Current grading schema foresees assessing three out of six quality characteristics of this model, namely functionality, efficiency and maintainability. The main goal of the Olympiad to emphasize problem solving, but not technical details of implementation, is a reasonable motivation why reliability and portability

are not assessed. As for usability, there are very concrete requirements for the data input/output and the programs, the compliance to these requirements may be checked many times during the contest by submitting the program to Contest and Grading System (and corrected if needed). The programs which do not obey the requirements are not forwarded for grading. The most important characteristic is functionality and other have meaning only if the program is functional. This (i.e. the level of functionality) is taken into account while assessing maintainability, but not when assessing efficiency. The most complicated is assessing functionality as there exists confrontation between three different points of views. From the point of view of software quality model, functionality is a Boolean value and there should be no partial score. From educational point of view, motivation of contestants has high priority in the contest, so partial scoring is very important, while from the point of view of problem solving, automated grading is not good enough as not equivalent algorithms may be assigned similar scores (i.e. testing reveals the existence of an error, but it doesn't show the scope of the error). It can be concluded that overall current grading schema in Lithuanian Olympiads corresponds ISO-9126-1 quality model, however, the assessment of functionality should be more investigated and motivated.