

# KONVERSATSIOONIAGENDI MODELLEERIMINE ARGUMENTEERIMISDIALOOGIS: SUHTLUS KUI INFOSEISUNDITE VÄRSKENDAMINE

Mare Koit

**Ülevaade.** Artiklis modelleeritakse sellist vestlusprotsessi, milles üks osaleja püüab mõjutada teist, et see otsustaks teha teatava tegevuse. Keskendutakse konversatsiooniagendi infoseisundite esitusele ja värskendamisreeglitele, mille rakendamise tulemusel liigub agent suhtluses ühest infoseisundist teise. Infoseisundi koosseisu kuulub partneri mudel, mis koosneb partneri oletatavatest hinnangutest vaadeldava tegevuse erinevatele aspektidele. Partneri mudel muutub suhtluse käigus, vastavalt argumentidele ja vastuargumentidele, mida osalejad esitavad tegevuse tegemise poolt või vastu. Konversatsiooniagendi mudeli praktilise realiseerimise kavandatakse dialoogsüsteemi, nn suhtlustreenerit. Artiklis arutatakse ka inimestevahelise suhtluse modelleerimise üle, tuues näiteid eestikeelsetest argivestlustest.\*

**Võtmesõnad:** argumenteerimisdialoog, dialoogi mudel, arutlusmudel, konversatsiooniagent, eesti keel

## 1. Sissejuhatus

Konversatsiooniagentide modelleerimise ja dialoogsüsteemide arendamise eesmärk on muuta inimeste ja arvuti vaheline suhtlus mugavamaks. Konversatsiooniagendid suhtlevad kasutajatega loomulikus keeles, inimestevahelise suhtluse reeglite alusel, aidates planeerida reise, vastates küsimustele näiteks ilma või spordivõistluste kohta, tegutsedes telefoniassistentidena või täites veelgi keerulisemaid ülesandeid (McTear 2004).

Igal dialoogsüsteemil on rida tüüpilisi komponente (Allen jt 2001). Kõnetuvastusmoodul leiab kasutaja sisendi tähenduse, tekstigenereerimis- ja kõnesünteesikomponent teisendavad dialoogsüsteemi väljundi tähenduse vastavalt tekstiks ja kõneks. Dialoogihaldur juhib konversatsiooniagendi käitumist, tehes selleks koostööd ülesannete halduriga, mis omakorda kasutab teadmisi ainevaldkonna kohta.

\* Artikli valmimist on toetanud Euroopa Regionaalarengufond Eesti Arvutiteaduse Tippkeskuse kaudu, Eesti Teadusfond (grandid 7503, 8558 ja 9124) ning Haridus- ja Teadusministeerium (projektid SF0180078s08, EKT5 ja EKT8). Autor tänab anonüümseid retsensente.

Dialoogihaldurites rakendatakse nelja põhilist ülesehitusviisi ehk arhitektuuri (Ginzburg, Fernandez 2010). Varaseim, aga samas praktiliseks realiseerimiseks kõige raskem arhitektuur pärineb 1960-ndate aastate tehisintellektist ja põhineb planeerimistehnikal (Allen 1995). Siin eeldatakse, et agendil on eesmärk/kavatsus, mis käivitab plaani koostamise selle eesmärgi saavutamiseks. Formaalselt võttes koosneb dialoog siis kahest otsinguprotsessist: planeerimine, kus rääkija kavatsustele seatakse vastavusse tegevused, ja plaani tuvastamine, kus kuulaja tuvastab rääkija kavatsused (Webber 2002). Plaanipõhised dialoogimudelid võtavad arvesse dialoogis osalejate eesmärke ja nende eesmärkide saavutamise viise. See võimaldab paindlikku suhtlust arvutiga. Kuid nagu juba öeldud, on selliste mudelite konstrueerimine ja arvutil rakendamine keeruline.

Järgmised kaks arhitektuuri, mis põhinevad vastavalt olekuautomaatidel ja freimidel, on kõige lihtsamad ja kommertsrakendustes kõige populaarsemad (Wilks jt 2005). Olemasolevad dialoogsüsteemid ongi enamasti realiseeritud kui olekuautomaadid, mis kasutavad regulaaravaldisi, nende hulgas on mitmesugused infosüsteemid, ärikokkulepete sõlmimise süsteemid, abilauad (ingl *help desk*). Nii saavutatakse praktilistes rakendustes vajalik tõrkekindlus, sest igas dialoogiolekus on kasutaja sõnavara ja muud valikud piiratud. Sel viisil saab lihtsustada dialoogi juhtimist ja kõnedialoogsüsteemides automaatset kõnetuvastust. Paraku jääb sellistel süsteemidel puudu paindlikkusest ja funktsionaalsusest, mis aga on inimestevahelise suhtluse olulised tunnused (Jurafsky, Martin 2008).

Neljas, kõige keerulisem ja võimsam dialoogihaldurite arhitektuur põhineb infoseisundi mõistel (Traum, Larsson 2003). Iga infoseisund tuleneb dialoogis eelnevatest seisunditest ja selles seisundis sooritatavad tegevused määravad omakorda järgmise infoseisundi. Dialoogihalduri funktsioone saab formaliseerida kui infoseisundite ajakohastamisi vastavalt etteantud nn värskendamisreeglitele. Dialoogi modelleerimisel on mõnikord vastandatud dialoogioleku mõistet kasutatavat lähenemisviisi (nt olekuautomaatides) plaanipõhisele lähenemisele. Infoseisundi mõiste võimaldab need kaks lähenemist ühendada, sest infoseisund võib sisaldada nii dialoogioleku komponente kui ka dialoogis osalejate arvamusi, soove, eesmärke jms, mida kasutab plaanipõhine lähenemine.

Tuleb veel märkida, et dialoogihaldurite ülesehitamisel on rakendatud lisaks eespool nimetatud reeglipõhiste meetoditele ka andmepõhiseid. Dialoogi on käsitletud kui Markovi otsustusprotsessi, kus on antud olekute hulk  $S = \{s_i | i=1, \dots, n\}$  ja võimalike tegevuste (dialoogiaktide) hulk  $A = \{a_j | j=1, \dots, m\}$ . Iga tegevusega on seotud "hüvitis", mis on seda suurem, mida edukam on tegevus lõppeesmärgi saavutamise seisukohast. Igas olekus valib dialoogsüsteem sellise tegevuse, mis maksimeerib tõenäolise hüvitise (Levin jt 2000). Selle mudeli nõrgaks küljeks on Markovi eeldus: iga järgmine olek loetakse sõltuvaks ainult vahetult eelnevast olekust ja täidetavast tegevusest (mis tegelikkuses nii ei tarvitse olla). Üldisem ja võimsam mudel on nn osaliselt vaadeldav Markovi otsustusprotsess (ingl *partially observable Markov decision process*), kus olekuteks on kasutaja oletatavad kavatsused (Roy jt 2000, Williams, Young 2007). Iga selline nn arvamusolek tuletatakse tõenäosusjaotusest, mis määratakse tegelike dialoogiandmete põhjal. Andmepõhiste meetodite rakendamist piirab aga empiirilise materjali kättesaadavus ja kogus.

Käesolevas artiklis vaadeldakse dialooge, kus ühe osaleja ( $A$ ) eesmärk on saavutada teise osaleja ( $B$ ) otsus teha tegevus  $D$ .  $A$  alustab dialoogi, tehes partnerile

teatavaks oma suhtluseesmärgi. Kui  $B$  keeldub (olles eelnevalt läbi viinud arutluse selle üle, kas  $D$  teha või tegemata jätta), siis peab  $A$  teda suhtluse käigus mõjutama, analüüsides selleks, millisel oma arutluse sammul jõudis  $B$  negatiivse otsuseni, ja samas korrigeerides olemasolevat partneri mudelit.

Artiklis arendatakse edasi mudelit, mida on kirjeldatud artiklis (Koit 2010). Artikli ülesehitus on järgmine. 2. osas antakse ülevaade suhtlusprotsessi modelleerimisest. Esitatakse konversatsiooniagendi mudel, mille koosseisu kuulub arutlusmudel. 3. osas vaadeldakse konversatsiooniagendi suhtlust partneriga kui protsessi, kus pidevalt värskendatakse infoseisundeid. 4. osas käsitletakse selle mudeli rakendamist dialoogsüsteemis, 5. osas vaadeldakse mudeli kasutatavust inimestevaheliste dialoogide analüüsimisel ning 6. osas tehakse kokkuvõtteid.

## 2. Suhtlusprotsessi modelleerimine

Toimugu suhtlus kahe osaleja – konversatsiooniagendi  $A$  ja tema partneri  $B$  vahel, kus  $B$  võib olla kas inimkasutaja või hoopis teine konversatsiooniagent.  $A$  eesmärgiks on mõjutada suhtluse käigus partnerit  $B$  vastu võtma otsust teha teatav tegevus  $D$ . Suhtlusprotsessi võib esitada järgmiste hulkade abil (Koit, Õim 2003):

- 1) suhtluseesmärkide hulk  $G$ , millest kumbki osaleja saab valida oma algse eesmärgi (vastavalt  $G^A$  ja  $G^B$ ). Antud juhul  $G^A = "B$  otsustab teha  $D"$ .
- 2) Osalejate suhtlusstrateegiate hulk  $S$ . Suhtlusstrateegia on algoritm, mida osaleja kasutab oma suhtluseesmärgi saavutamiseks. Strateegia määrab osaleja tegevuse suhtluse igal sammul.
- 3) Suhtlustaktikate hulk  $T$ . Suhtlustaktika on osaleja järjekordse vooor konstrueerimise algoritm, mis määrab suhtluspartneri mõjutamise viisi. Näiteks võib  $A$  oma suhtluseesmärgi  $G^A$  saavutamiseks ahvatleda, veenda või ähvardada partnerit tegema  $D$ , rõhutades selle tegevuse erinevaid aspekte.
- 4) Arutlusmudelite hulk  $R$ , mida osalejad kasutavad, arutlemaks tegevuse  $D$  tegemise või mittetegemise üle. Arutlusmudel on algoritm, mille tulemuseks on kas positiivne või negatiivne otsus arutelu objekti kohta (antud juhul on selleks tegevuse  $D$  tegemine).
- 5) Osalejate mudelite hulk  $P$ . Osaleja mudel on osaleja ettekujutus kas iseendast või oma partnerist:  
$$P = \{P^A(A), P^A(B), P^B(A), P^B(B)\}.$$
- 6) Keeleteadmiste hulk, mida osalejad kasutavad oma lausungite sünteesimisel ja partneri lausungite analüüsimisel.
- 7) Maailmateadmiste hulk, mida osalejad kasutavad arutlusprotsessis.

### 2.1. Arutlusmudel

Arutlusprotsess, kus tuleb vastu võtta otsus tegevuse  $D$  tegemise kohta, kujutab endast kindlate sammude järjendit, kus kaalutakse tegevuse tegemiseks vajaminevaid/olemasolevaid ressursse ning tegevuse positiivseid ja negatiivseid aspekte. Tegevuse positiivsed aspektid on näiteks tema meeldivus ja kasulikkus, negatiivsed näiteks ebameeldivus ja kahjulikkus. Suhtluspartner ei saa otseselt osaleda selles

arutluses, sest see toimub teise osaleja peas. Ainus, mida ta saab teha, on oma lausungite abil püüda juhtida partneri arutlusprotsessi, andes (uut) infot tegevuse aspektide kohta ning püüdes rõhutada positiivseid ja alla suruda negatiivseid aspekte.

Siin kasutatav arutlusmudel koosneb kahest osast: 1) inimese motivatsioonisfääri mudelist ja 2) arutlusprotseduuridest. Subjekti motivatsioonisfääri mudelina kasutatakse vektorit, mille komponentideks on “kaalud”, mida subjekt on oletatavasti omistanud arutlusobjektiks oleva tegevuse aspektidele:

$w = (w(\text{ressursid}), w(\text{meeldiv}), w(\text{ebameeldiv}), w(\text{kasulik}), w(\text{kahjulik}), w(\text{kohustuslik}), w(\text{keelatud}), w(\text{karistus-keelatud-tegevuse-tegemise-eest}), w(\text{karistus-kohustusliku-tegevuse-tegematajätmise-eest}))$ .

Siin tähistab  $w(\text{ressursid})$  ressursside olemasolu/puudumist (vastavalt 1 või 0) ning  $w(\text{meeldiv})$  jne  $D$  meeldivuse jt aspektide kaalu. Vektorit  $w^{AB}$  kasutab  $A$  kui partneri mudelit  $P^A(B)$ , s.t see on tema ettekujutus partnerist. Tegevuse  $D$  aspektide kaalud on  $A$  arvamused/oletused  $B$  hinnangute kohta (mille tegelikke väärtusi ta ei tarvitse teada). Suhtluse kestel teeb  $A$  vastavalt vajadusele korrekture partneri mudelis.

Arutlusmudeli teine osa on arutluskkeemid, mis oletatavasti reguleerivad inimese arutlusi tegevuste üle. Arutluskkeem kujutab endast nn arutluskammude järjendit, mida arutlev subjekt läbib arutlusprotsessis. Iga samm seisneb teatavate kaalude võrdlemises ja lõpptulemuseks on otsus: teha  $D$  või mitte teha  $D$  (vt Koit, Õim 2003). Motivatsioonisfääris saab eristada kolme olulist tegurit, mis reguleerivad subjekti arutlust tegevuse  $D$  üle. Esmalt, subjekt soovib teha  $D$ , kui  $D$  meeldivus ületab ebameeldivust. Teiseks, subjektil on vaja teha  $D$ , kui kasulikkus ületab kahjulikkust. Kolmandaks, subjekt peab (on sunnitud) tegema  $D$ , kui see on talle kohustuslik, s.t kui  $D$  tegematajätmine tooks kaasa karistuse. Neid tegureid võib nimetada vastavalt soov-, vaja- ja peab-teguriteks. Igauks neist käivitab teatava arutlusprotseduuri, olgu need protseduurid vastavalt SOOV, VAJA ja PEAB.

Siin eeldatakse, et vektori  $w$  kõikidel komponentidel on arvulised väärtused ja arutlusprotsessis, kus kaalutakse argumente ja vastuargumente, liidetakse ja võrreldakse neid kaalusid. (Muidugi on see robustne eeldus, mis lihtsustab tegelikkust.)

Üldjoontes järgib see arutlusmudel BDI-mudeli (ingl *belief-desire-intention*) ideid (Allen 1995).

## 2.2. Arutlus suhtluses

Ühe osaleja (konversatsioonigendi  $A$ ) eesmärkide baasis aktiveeritakse eesmärk  $G^A$ , mis on seotud partneri  $B$  tegevustega.  $A$  kontrollib partneri mudelit, mis koosneb  $D$  aspektide oletatavatest kaaludest, mida  $B$  on  $A$  arvates neile aspektidele omistanud. Selle mudeli alusel valib  $A$  taktika, mida kasutada  $B$  mõjutamisel (näiteks veenda  $B$ -d tegema  $D$ , s.t rõhutada  $D$  tegemise kasulikkust  $B$  jaoks). Seega püstitab  $A$  alameesmärgi – käivitada  $B$ -s kindel arutlusprotseduur (näiteks veenmise puhul käivitada arutlusprotseduur VAJA, mis lähtub vaja-tegurist). Konstrueerides oma esimest vooru  $vr_1$ , kavandab  $A$  selles voorus esinevad dialoogiaktid ja määrab nende keelelise vormi.  $A$  voor  $vr_1$  käivitab  $B$ -s arutlusprotsessi, kus ta kõigepealt interpreteerib  $A$  vooru ja seejärel genereerib oma vooru  $vr_2$ .  $B$  vastus käivitab omakorda

$A$ -s arutluse, mille käigus  $A$  kontrollib, kuidas on kulgenud tema suhtluseesmärgi saavutamine, ja vastavalt aktiveerib  $G^A$  uue alameesmärgi. Tsükkel kordub:  $A$  ehitab oma järgmise vooru  $vr_3$  jne. Dialoog lõpeb, kui  $A$  on kas saavutanud oma eesmärgi või sellest loobunud.

### 3. Suhtlus kui infoseisundite värskendamine

#### 3.1. Infoseisundite esitus

Infoseisundi oluliseks komponendiks on partneri mudel, mis suhtluse käigus pidevalt muutub.

Konversatsiooniagendi infoseisund koosneb kahest osast – privaatsest ja jagatud osast. Privaatne osa sisaldab infot, mis on kättesaadav ainult konversatsiooniagendile, seevastu jagatud osas olev info on kättesaadav ka partnerile.

Siin tehakse veel üks lihtsustav eeldus: konversatsiooniagendi jaoks on ette antud valmis laused, mille hulgas ta saab valida argumente partneri mõjutamiseks. Need laused on klassifitseeritud vastavalt vaatluse all oleva tegevuse  $D$  aspektidele ning nende aspektide kaalude muutmiseks kasutatavatele dialoogiaktidele (nt meeldivuse suurendamiseks, ebameeldivuse vähendamiseks jne). Seega lihtsustub konversatsiooniagendi voorude genereerimine: tal on vaja üksnes valida laused õigetest klassidest.

(Jooksva) infoseisundi privaatne osa võib sisaldada näiteks järgmised komponendid:

- jooksev partneri mudel (kaalude vektor  $w^{AB}$  –  $A$  ettekujutus  $B$  hinnangutest tegevusele  $D$ ). Suhtluse algul võib  $A$  konstrueerida selle vektori kas juhuslikult või – kui võimalik – võtta arvesse oma eelteadmisi  $B$  kohta.
- Suhtlustaktika  $t_i^A$ , mille  $A$  on valinud  $B$  mõjutamiseks (lähtuvalt partneri mudelist).
- Arutlusprotseduur  $r_j$ , mida  $A$  püüab partneris  $B$  käivitada ja seejärel juhtida positiivse otsuseni (see on määratud  $A$  taktikaga, nt kui  $A$  on valinud veenmise taktika, siis püüab ta  $B$ -s käivitada arutlusprotseduuri VAJA).
- (Alam-)eesmärkide pinu (magasin), mis sisaldab püstitatud, kuid seni veel saavutamata eesmärgid. Alguses paigutab  $A$  pinusse oma lähte-eesmärgi (“ $B$  otsustab teha  $D$ ”). Igas infoseisundis sisaldab pinu kindlasti ka  $D$  selle aspekti, mis on jooksva taktika nn nimiaspekt. (Näiteks kui  $A$  alustab  $B$  veenmist, siis on pinu tipus eesmärk “kasu”: suurendada  $D$  kasulikkust  $B$  jaoks.)
- Võimalike dialoogiaktide hulk  $DA = \{d_1^A, d_2^A, \dots, d_n^A\}$ .  $A$  kui suhtluse algataja saab kasutada näiteks järgmisi dialoogiakte: ettepanek, väited  $D$  erinevate aspektide kaalude suurendamiseks ja vähendamiseks jne.
- (Lõplik) hulk lausungeid, mida  $A$  saab kasutada tegevuse  $D$  aspektide kaalude suurendamiseks/vähendamiseks (s.t argumendid ja vastuargumendid  $D$  tegemisele):  $U = \{u_{i1}^A, u_{i2}^A, \dots, u_{iki}^A\}$ . Igale sellisele lausungile olgu samuti omistatud kaal (arvuline väärtus) nagu tegevuse  $D$  aspektidele, mis tähendab, et mõned argumendid on suurema ja teised väiksema kaaluga:  $V = \{v_{i1}^A, v_{i2}^A, \dots, v_{iki}^A\}$ , kus  $v_{i1}^A$  jne on vastavalt  $u_{i1}^A$  jne väärtus. Siin

eeldatakse, et  $A$  saab iga lausungit valida üheainsa korra. See tähendab, et kui argumentid lõpevad, siis peab  $A$  oma eesmärgist loobuma.

Infoseisundi jagatud osa sisaldab näiteks järgmised komponendid:

- arutlusmudelite hulk  $R = \{r_1, \dots, r_k\}$ ,
- suhtlustaktikate hulk  $T = \{t_1, t_2, \dots, t_p\}$ ,
- dialoogi ajalugu – dialoogis (seni) esinenud lausungite järjend koos osalejate tähistega ja vastavate dialoogiaktidega:

$$p_1:u_1[d_1], p_2:u_2[d_2], \dots, P_i:u_i[d_i],$$

kus  $p_1, p_2$  jne on osaleja tähis  $A$  või  $B$ .

### 3.2. Värskendamisreeglid

Dialoogi käigus toimub ühest infoseisundist teise liikumine nn värskendamisreeglite abil, mida saab jagada kahte liiki: genereerimis- ja interpreteerimisreeglid.

I. Reeglid, mida  $A$  kasutab oma voorude genereerimiseks, saab omakorda liigitada vastavalt rakendatavale taktikale ja suhtlussituatsioonile. Siia kuuluvad reeglid erinevate juhtude käsitlemiseks, näiteks:

- 1) juhul, kui jooksva taktika “nimiaspekt” paikneb parajasti pinu tipus (näiteks kui jooksev taktika on veenmine, siis on nimiaspektiks “kasu”).
- 2) Juhul, kui nimiaspekti peal paikneb pinus mõni muu aspekt. (Näiteks kui  $A$  püüab  $B$ -d veenda, rõhutades kasu, aga  $B$  viitab hoopis tegevuse ebameeldivusele, siis asub “ebameeldivus” pinus “kasu” peal, sest  $A$  peab kõigepealt püüdma vähendada tegevuse ebameeldivust, mille tõi esile  $B$ , ja alles seejärel saab suurendada kasulikkust.)
- 3) Juhul, kui jooksva taktika jätkamiseks puuduvad lausungid (ja  $A$  peab valima uue taktika, kui see on võimalik).
- 4) Juhul, kui  $A$  peab loobuma oma algsest eesmärgist.
- 5) Juhul, kui  $B$  on teinud positiivse otsuse ja  $A$  on seega saavutanud oma eesmärgi.

Suhtluse alustamiseks, s.t esmase infoseisundi värskendamiseks, on omaette reeglid.

II. Reeglid, mida  $A$  kasutab  $B$  lausungite interpreteerimiseks. Nende reeglite alusel saab  $A$  tuvastada  $B$  lausungi sisu (näiteks tegevuse kahjulikkuse osutamine) ja kasutatud dialoogiakti (näiteks keeldumine).

Kahest järgnevast värskendamisreegli näitest kuulub (1) liiki II ja (2) liiki I-2.

- (1) kui jooksev taktika on veenmine JA  $B$  viimane lausung käsitles  $D$  ebameeldivust, SIIS paiguta eesmärkide pinusse esmalt “kasu” ja seejärel “ebameeldivus”.
- (2) kui jooksev taktika on veenmine JA eesmärkide pinu tipus on “ebameeldivus” (s.t  $B$  ütles, et  $D$  tegemine on ebameeldiv) JA “kasu” paikneb eesmärkide pinu tipus oleva kaalu all JA leidub lausungeid ebameeldivuse vähendamiseks  $x$  ühiku võrra JA leidub lausungeid kasulikkuse suurendamiseks  $y$  ühiku võrra JA arutlusprotseduur VAJA korrigeeritud partneri mudelil (kus  $w(\text{ebameeldivus})$  on  $x$  ühiku võrra väiksem kui seni ja  $w(\text{kasulikkus})$

on  $y$  ühiku võrra suurem kui seni) lõpeb otsusega “teha  $D$ ”, siis vali need lausungid (ja vastavad dialoogiaktid)  $J_A$  eemalda eesmärkide pinust “ebameeldivus” ning “kasu”.

Reegleid (2) ja (1) rakendatakse juhul, kui agent  $A$  on valinud partneri  $B$  veenmise taktika, mille puhul ta rõhutab tegevuse  $D$  kasulikkust, aga partner – seadmata küll kahtluse alla kasulikkust – viitab oma järjekordses voorus hoopis tegevuse ebameeldivusele. Agent peab reageerima partneri lausungile, näidates, et tegevus ei ole nii ebameeldiv, nagu arvab partner, aga – järgides valitud taktikat – ühtlasi rõhutama tegevuse kasulikkust.

Eesmärk eemaldatakse pinust niipea, kui agent leiab, et arutlus partneri mudeli järgi annab positiivse otsuse. Kui partneri järgmisest lausungist selgub, et eesmärk siiski ei ole saavutatud, siis pannakse see pinusse tagasi ja protsess jätkub.

## 4. Arutelu

Kui  $A$  püüab mõjutada  $B$  otsuseni jõudmist, siis kasutab ta mitmesuguseid väiteid käsitletava tegevuse  $D$  positiivsete aspektide kaalude suurendamiseks ja negatiivsete aspektide kaalude vähendamiseks, et tulemuseks oleks  $B$  otsus teha  $D$ .

Kui  $B$  viitab oma lausungis kindlale aspektile, mille tegelik kaal (liiga väike või liiga suur) ei võimalda tal vastu võtta otsust teha  $D$ , siis saab  $A$  lihtsalt valida ühe võimalikest argumentidest selle aspekti “tõrjumiseks”. Näiteks kui  $B$  ütleb, et tegevuse  $D$  tegemiseks puuduvad tal ressursid, siis saab  $A$  valida lausungi, mis näitab, et ressursid on hangitavad ja kuidas see saaks toimuda. Kui  $B$  ei põhjenda oma negatiivset otsust, siis saab  $A$  üksnes rõhutada kasutatava taktika nimiaspekti, näiteks veenmise korral kasu.

Dialoogi alustades moodustab konversatsiooniagent partneri mudeli  $w^{AB}$  (kaalude vektori) ja valib taktika, mis sellel mudelil viiks positiivse otsuseni. Teiste sõnadega,  $A$  eeldab, et  $B$  on valitud taktika abil mõjutatav. Dialoog algab sellega, et  $A$  väljendab oma suhtluseesmärki (näiteks teeb  $B$ -le ettepaneku teha tegevus  $D$ ). See on agendi esimene voor  $vr_1$ . Kui kasutaja keeldub (olles kaalunud tegevuse aspektide tegelikke, temal olemasolevaid väärtusi), siis leiab agent kasutaja lausungist  $vr_2$  tegevuse  $D$  selle aspekti, mille kaalu tuleks muuta, et kasutaja mudelit tegelikusega kooskõlla viia. Siin eeldatakse, et igal lausungil (argumendil) on arvuline väärtus ja selle lausungi ütlemine muudab vastavat kaalu partneri mudelis. Oma vooru  $vr_3$  konstrueerides valib arvuti olemasolevate lausungite hulgast lausungi selle kaalu muutmiseks ja ühtlasi muudab sellele lausungile omistatud väärtuse võrra vastavat kaalu mudelis  $w^{AB}$ . Eeldatavasti annab arutlusprotseduur muudetud mudelil positiivse tulemuse (teha  $D$ ), sest  $A$  peaks valima just sellise lausungi, et tulemuseks oleks  $B$  positiivne otsus. Nüüd on kasutaja kord valida vastuslausung  $vr_4$  ja protsess jätkub analoogilisel viisil.

Järgnevas lihtsas näites (3) on tegevuseks  $D$  “valmistada kartulisalat”:  $A$  eesmärk on, et  $B$  otsustaks teha  $D$  (vrd Koit 2011).  $A$  genereerib sellise partneri mudeli, millel arutlusprotseduur  $VAJA$  annab positiivse tulemuse,  $A$  rakendab veenmise taktikat.

Olgu konversatsiooniagendi  $A$  algne (esmane) infoseisund järgmine.

Privaatne osa:

- algne partneri mudel  
 $w^{AB} = (w^{AB}(\text{ressursid}) = 1, w^{AB}(\text{meeldiv}) = 3, w^{AB}(\text{ebameeldiv}) = 0, w^{AB}(\text{kasu}) = 7, w^{AB}(\text{kahju}) = 0, w^{AB}(\text{kohustuslik}) = 0, w^{AB}(\text{keelatud}) = 0, w^{AB}(\text{karistus-keelatud-tegevuse-tegemise-est}) = 0, w^{AB}(\text{karistus-kohustusliku-tegevuse-tegematajätmise-est}) = 0).$
- *A* valitud taktika – veenmine.
- *A* püüab käivitada *B* arutlusprotseduuri VAJA (kuna selle eeltingimused on täidetud:  $w^{AB}(\text{ressursid}) = 1; w^{AB}(\text{kasu}) > w^{AB}(\text{kahju})$ ).
- Eesmärkide pinu sisaldab ainult *A* algse eesmärgi (“*B* otsustab teha *D*”).
- *A* käsutuses olevate dialoogiaktide loend: {ettepanek, väited *D* erinevate aspektide kaalude suurendamiseks/vähendamiseks jne}.
- Lausungite hulk erinevate dialoogiaktide väljendamiseks (sh kaalude suurendamiseks/vähendamiseks, koos nende väärtustega): {*Ma aitan sind* – väärtus 1, *Minu köögis on hea ventilatsioon* – väärtus 5 jne}.

Esmase infoseisundi jagatud osa sisaldab:

- arutlusprotseduurid: SOOV, VAJA, PEAB.
- Taktikad: ahvatlemine, veenmine, ähvardamine.
- Dialoogi ajalugu on tühi hulk.

*A* esimene dialoogiakt on ettepanek, mida väljendab lausung *Palun valmista kartulisalat* (3):

(3)

*A* (arvuti): Palun valmista kartulisalat. [Ettepanek.]

*B* (kasutaja): Mul pole nii palju aega. [Keeldumine *D* tegemisest + väide, mis vähendab aspekti ressursi väärtust.]

Seega on aspekti  $w^B(\text{ressursid})$  tegelik väärtus 0 ja arvuti peab korrigeerima komponendi  $w^{AB}(\text{ressursid})$  väärtust partneri mudelis  $w^{AB}$ . Seepärast valib ta väite selle väärtuse suurendamiseks (ehk argumendi *D* tegemise kasuks):

*A*: Ma aitan sind. [Argumendiga mittenõustumine + väide komponendi  $w^{AB}(\text{ressursid})$  kaalu suurendamiseks: koos tegutsemine võimaldab aega kokku hoida.]

*B*: Köögis on liiga palav. [Keeldumine *D* tegemisest + argumendiga mittenõustumine + väide, mis suurendab aspekti kahju väärtust.]

Kasutaja tõi esile tegevuse *D* kahjulikkuse. Järelikult peab arvuti korrigeerima komponendi  $w^{AB}(\text{kahju})$  kaalu kasutaja mudelis:

*A*: Minu köögis on hea ventilatsioon. [Argumendiga mittenõustumine + väide aspekti kahju kaalu vähendamiseks.]  
jne.

Tartu Ülikoolis on arendamisel dialoogsüsteem, mis kasutajaga suhtlemisel võib täita kas *A* või *B* rolli. Hetkel opereerib arvuti ainult keelelise sisendi ja väljundi semantiliste esitustega; nii arvuti kui ka kasutaja jaoks võimalike eelnevalt klassifitseeritud lausungite hulgad on ette antud. Sellisest programmist võiks olla abi argumenteerimise harjutamisel. Kui kasutaja täidab osaleja *A* rolli, siis võib ta püüda mõjutada oma “vastasmängijat” *B* (arvutit) näiteks sel teel, et esitab süstemaatiliselt ainult meeldivust rõhutavaid ja ebameeldivust allasuruvaid argumente, s.t rakendab



ahvatlemise taktikat ega “hüple” ühelt taktikalt teisele. Samas, kui kasutaja mängib *B* rolli ja arvuti rakendab järjekindlalt ühte fikseeritud taktikat, saab kasutaja harjutada kindlate argumentide tõrjumist kindlate vastuargumentidega. Sel viisil saab distsiplineerida suhtlemist ja argumentide esitamist oma seisukohtade kaitseks.

## 5. Kuidas inimesed argumenteerivad?

Kas eespool toodud mudel võimaldab kirjeldada ka tegelikke, inimeste vahel toimuvaid dialooge? Käesoleva artikli tarvis analüüsiti Tartu Ülikooli eesti dialoogikorpusest üheksat argivistlust, kus leidis argumenteerimist. Argivistlustes on tavaline, et arendatakse korruga mitut teemat, mistõttu jutt liigub siia-sinna ning järjekindlat argumenteerimist kohtab harva. Siiski on analüüsitud dialoogides lõike, kus kõne all on teatava tegevuse tegemine ning sellele tuuakse argumente ja vastuargumente.

Järgnevas näites (4) püüab osaleja *A* saavutada partneri *B* nõusolekut, et too paigutaks oma nime ja muud andmed ühe instituudi veebilehele. *B* kuulub teatava nõukogu juhatusse, kuid ei pea vajalikuks oma andmete avalikku seostamist selle instituudiga. Dialoogide litereerimisel on kasutatud vestlusanalüüsi transkriptsiooni, dialoogiaktid on märgendatud TÜ aktitüpoloogia kohaselt (Hennoste, Rääbis 2004).

(4)

*B*: a:ga ma mõtlen oopis `teist asja, | YA: EELTEADE |  
[võib]olla ma=ei:: `pane sinna `ültse mingeid andmeid. | SEE: ARVAMUS |  
[Keeldumine.]

/---/

*A*: aga [kõik `teised] `panevad. | SEJ: MITTENÕUSTUMINE | | SEE: VÄIDE | [Argumentidiga mittenõustumine + väide aspekti kasu kaalu suurendamiseks: kasulik on olla samasugune nagu teised.]

/---/

*B*: jaahhhhh ((ohates)) | SEJ: NÕUSTUMINE |  
ei=no ma=i ma=i `näe nagu suurt `vajadust `miks: `miks see peaks seal isegi `olema. | SEE: ARVAMUS |  
[Keeldumine *D* tegemisest + väide, mis vähendab aspekti kasu väärtust.]  
(0.3) ausalt.=hh= | IL: MUU |

*A*: =noh ma=i=tea esiteks on see .hhhhhh nagu nimodi=hhhh .hhhhh  
`otsustatud=et=hh et nõukogu `tuleb=ja:=ja q noh (1.0) .hhhhhhhhh  
eeee ikkagi noh, see on ka `enda tutvustamise=hhhhhhh `mõttes kogu=se  
.hhhhhh > Kaerajaani instituudi kodu`lehekülg.=hhh < .hhh | SEJ:  
MITTENÕUSTUMINE | | SEE: VÄIDE | [Väide aspekti kasu kaalu suurendamiseks:  
kasu pole isiklik, vaid institutsionaalne.]

*B*: ei seda `küll= | SEJ: PIIRATUD NÕUSTUMINE |  
aga ma mõtlen mõtlen noh igasuguseit meili`aadresse=ja `vär[ke=ja]  
| IL: TÄPSUSTAMINE | | SEE: ARVAMUS | [Keeldumine *D* tegemisest + väide, mis suurendab aspekti kahju väärtust: avalikustatakse personaalseid andmeid.]  
jne.

Osalejate lausungid moodustavad seisukohavõtude naabruspaare (SEE ja SEJ), kusjuures iga seisukohavõtt kujutab endast ühtlasi argumenti või vastuargumenti.

*A* rakendab vestluses veenmise taktikat, rõhutades tegevuse kasulikkust. *B*, püüdes edasi lükata otsuse vastuvõtmist, viitab enamasti (väikesele) kasule või (suurele) kahjule. Vestluse lõpuks ei teegi *B* selget otsust.

Analüüsitud dialoogide põhiosa võib kujutada (pärast mõnede dialoogiaktide, nt IL, YA jmt, “destilleerimist”, vt Jönsson, Dahlbäck 2000) enamasti järgmise dialoogiaktide järjendina (5). Lausungid on varustatud kahe märgendiga, moodustades “kokkusulavaid” naabruspaare (eelneva paari järelliige, mis väljendab mittenõusolekut, on samal ajal järgneva paari esiliige). See näitab muuhulgas, et osalejate suhtluseesmärgid on vähemalt suhtluse algul vastandlikud.

(5)

*A*: DIE: ETTEPANEK / SOOV

*B*: DIJ: MITTENÕUSTUMINE / PIIRATUD NÕUSTUMINE + SEE: ARVAMUS / VÄIDE

*A*: SEJ: MITTENÕUSTUMINE + SEE: ARVAMUS / VÄIDE

*B*: SEJ: MITTENÕUSTUMINE / PIIRATUD NÕUSTUMINE + SEE: ARVAMUS / VÄIDE

...

*A*: SEJ: MITTENÕUSTUMINE + SEE: ARVAMUS / VÄIDE

*B*: SEJ: NÕUSTUMINE / PIIRATUD NÕUSTUMINE / MITTENÕUSTUMINE + DIJ: NÕUSTUMINE / PIIRATUD NÕUSTUMINE / MITTENÕUSTUMINE

Missuguseid taktikaid *A* ja *B* rakendavad ning kui järjekindlalt, see sõltub konkreetsete arvamuste või väidete sisust. Näiteks *A* taktika võib sellise dialoogiaktide järjendi puhul olla nii ahvatlemine, veenmine ja ähvardamine kui ka mõni selline, mida eespool toodud mudelis pole otseselt käsitletud.

Järgnev näide (6) illustreeribki ühte omapärast taktikat: *A*, kes soovib, et *B* keedaks talle kas piparmündi- või pärnaõieteed (see on siin tegevus *D*), kordab kogu vestluse jooksul ühtainsat (vastu-)argumenti (et mustsõstratee, mida kavatseb keeta *B*, on liiga mõru, s.t ebameeldiv), kuid lõpuks nõustub siiski *B*-ga. Niisiis jääb *A*-l saavutamata esialgne eesmärk, aga veelgi enam – ta nõustub *B*-ga, et see, mida ta algul pidas ebameeldivaks, on tegelikult hoopis meeldiv. Rollid justkui vahetuvad: *A* loovutab initsiatiivi *B*-le.

(6)

*A*: `tee mulle [ `ole=ea] | DIE: SOOV |

*B*: [ `mustsõ]strateed.= | DIE: PAKKUMINE | [*B* soovib tegevuse *D* asemel teha tegevust *D*.]

/---/

*A*: `piparmünt või [seda] | DIE: SOOV |

{*B*}: [hehe] | YA: MUU |

(0.5)

*A*: [ `pärnaõit.] | PA: ENESEPARANDUS | | DIE: SOOV |

/---/

*A*: ta=on `minu jaoks on `liiga [siuke] `mõru. | SEE: VÄIDE | [*A* argument: *D*' tulemus – mustsõstratee – on mõru, s.t ebameeldiv.]

*B*: % [ei= `ole `mõru.] % | SEJ: MITTENÕUSTUMINE | | SEE: VÄIDE | [*B* vastuargument: lihtsalt eitab *A* esitatud omadust, ei põhjenda. Sellised “argumendid” on inimestevahelises suhtluses võimalikud.]

(0.4)

A: mulle maitseb `jubedalt `mustsõstar aga `mustsõstra (.) [see] `tee on noh \$ siukse pagana `mõruda | SEJ: PIIRATUD NÕUSTUMINE | | SEE: VÄIDE | [A kordab oma argumenti: *D'* tulemus on ebameeldiv.]

{B}: [heheh] | YA: MUU |

(0.5)

{B}: \$ .nhh \$ | YA: MUU |

(1.1)

A: \$ siukse: \$ | YA: MUU |

{B}: heh | YA: MUU |

(1.7)

A: no igal=jul `mõruda. | IL: ÜLERÕHUTAMINE | [A kordab oma argumenti: *D'* tulemus on ebameeldiv.]

/---/

A: [{-}] mina tahan öelda \$ mustsõstratee on `mõru onju i- ikka \$ | SEE: VÄIDE | [A kordab oma argumenti: *D'* tulemus on ebameeldiv.]

{-}: he[hehe] | YA: MUU |

A: \$ [ja siis] ja siis tulevadki sellised et @ mm. @ \$ | SEE: VÄIDE |

(.)

B: väga `ea=jü sis ongi: | SEE: ARVAMUS | [Huvitav olukord: *B* väidab, et *A* esiletoodud negatiivne omadus on tegelikult positiivne (mõru tee ei ole ebameeldiv, vaid hoopis meeldiv).]

A: \$ jah \$ | SEJ: NÕUSTUMINE | (0.4) [Ja *B* nõustub sellega!]

Seega õnnestus hoopis *B*-l saavutada *A* nõusolek, et *B* teeb tegevuse *D* asemel tegevuse *D'*. *B* ei teinud selleks mitte midagi muud, kui ignoreeris korduvalt *A* argumenti (et *D'* tulemus on ebameeldiv) ja viimaks väitis, et see, mida *A* peab ebameeldivaks, on tegelikult meeldiv (“ongi ju hea”). Sellega *A* nõustus; sama naabruspaari SEE: ARVAMUS ja SEJ: NÕUSTUMINE korratakse veel kaks korda. Seega jõuavad *A* ja *B* vestluse lõpuks kokkuleppele, et *B* keedab mustsõstrateed.

Kõigis analüüsitud argivistlustes on suhtlejad omavahel tuttavad või koguni sõbrad, niisiis teavad teineteisest küllaltki palju. See lihtsustab taktikate valimist oma suhtluseesmärkide saavutamiseks, sest ettekujutus partnerist (partneri mudel) vastab üsna täpselt tegelikkusele ning ühest infoeisundist teise liikumine on seetõttu sujuv.

Tuleb siiski tõdeda, et reaalsed inimestevahelised dialoogid on palju rikkalikumad kui need, mida suudaks läbi viia eespool modelleeritud konversatsiooniagent isegi siis, kui lisada talle keele- ja maailmateadmised.

## 6. Kokkuvõte

Artiklis vaadeldi dialooge, kus ühe osaleja (*A*) eesmärk on, et partner (*B*) võtaks vastu otsuse teha tegevus *D*. Seda tüüpi dialoogid kuuluvad argumenteerimisdialoogide hulka. Kui *A* on vestluse algul väljendanud oma eesmärki (teinud ettepaneku), saab *B* vastata kas nõustumise või keeldumisega, sõltuvalt oma arutluse tulemusest. *B* võib põhjendada keeldumist, lisades sellele argumendi(d). *B* argumendid

annavad *A*-le infot *B* arutlusprotsessi kohta ning *A* saab edasise dialoogi käigus neid oma vastuargumentidega tõrjuda.

Artikkel kirjeldas arvutusmudelit, mida rakendatakse eksperimentaalses dialoogsüsteemis. Keskenduti konversatsiooniagendi infoseisundite esitusele ja värskendamisreeglitele. Infoseisund sisaldab partneri mudeli, mis koosneb tema oletatavatest hinnangutest kõnealuse tegevuse erinevatele aspektidele. Suhtluse käigus muutub partneri mudel pidevalt, vastavalt esitatavatele argumentidele ja vastuargumentidele. Praktilise rakendusena peetakse silmas nn suhtlustreenerit – programmi, millega kasutaja saab harjutada argumenteerimist.

Edasise töö suunad on järgmised. Detailiseerida mudelit, vaadeldes erinevaid stsenaariume: kui *A* ja *B* eesmärgid on vastandlikud ja üks osalejaist peab oma algsest eesmärgist loobuma (see on seni vaadeldud stsenaarium); kui *A* ja *B* teevad koostööd ühise eesmärgi nimel; kui *A* ja *B* on mõlemad konversatsiooniagendid, kumbki oma infoseisundite ja värskendamisreeglitega. Analüüsida erinevaid suhtlusstrateegiaid ja -taktikaid, võttes arvesse nende edukust osalejate algsete eesmärkide saavutamisel.

### Viidatud kirjandus

- Allen, James 1995. Natural Language Understanding. Second Edition. Redwood City etc.: The Benjamins/Cummings Publ. Co.
- Allen, James; Ferguson, George; Stent, Amanda 2001. An architecture for more realistic conversational systems. – IUI'01: Proceedings of the 6th International Conference on Intelligent User Interfaces. Santa Fe, New Mexico, United States: ACM Press, 1–8. <http://dx.doi.org/10.1145/359784.359822>
- Ginzburg, Jonathan; Fernández, Raquel 2010. Computational models of dialogue. – A. Clark, C. Fox, S. Lappin (Eds.). The Handbook of Computational Linguistics and Natural Language Processing. Oxford: Wiley-Blackwell, 429–481. <http://dx.doi.org/10.1002/9781444324044.ch16>
- Hennoste, Tiit; Rääbis, Andriela 2004. Dialoogiaktid eesti infodialoogides: tüpoloogia ja analüüs. Tartu: Tartu Ülikooli Kirjastus.
- Jurafsky, Daniel; Martin, James H. 2008. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Prentice Hall.
- Jönsson, Arne; Dahlbäck, Nils 2000. Distilling dialogues – A method using natural dialogue corpora for dialogue systems development. – Proceedings of the 6th Applied Natural Language Processing Conference. Seattle, 44–51.
- Koit, Mare 2011. Conversational agent in argumentation: Updating of information states. – J. Filipe, J. L. G. Dietz (Eds.). Proceedings of the International Conference on Knowledge Engineering and Ontology Development: KEOD 2011, Paris, France, 26–29 October, 2011. Paris: SciTEC Publications Ltd, 375–378.
- Koit, Mare 2010. Eesti dialoogikorpuse ja argumenteerimisdialogi arvutil modelleerimine. – Keel ja Kirjandus, 4, 241–262.
- Koit, Mare; Õim, Haldur 2003. Eestikeelse dialoogi modelleerimine. – Keel ja Kirjandus, 10, 721–735.
- Levin, Esther; Pieraccini, Roberto; Eckert, Wieland 2000. A Stochastic Model of Human-Machine Interaction for Learning Dialogue Strategies. – IEEE Transactions on Speech and Audio Processing, 8 (1), 11–23. <http://dx.doi.org/10.1109/89.817450>
- McTear, Michael 2004. Spoken Dialogue Technology. Toward the Conversational User Interface. London: Springer.

- Roy, Nicholas; Pineau, Joelle; Thrun, Sebastian 2000. Spoken Dialogue Management Using Probabilistic Reasoning. – Annual meeting of the Association for Computational Linguistics (ACL-2000).
- Traum, David; Larsson, Staffan 2003. The Information State Approach to Dialogue Management. – Jan van Kuppevelt, Ronnie W. Smith (Eds.). Current and New Directions in Discourse and Dialogue. Dordrecht: Kluwer, 325–353.
- Webber, Bonnie L. 2002. Computational perspectives on discourse and dialogue. – D. Schiffrin, D. Tannen, H. Hamilton (Eds.). The Handbook of Discourse Analysis. Oxford: Blackwell Publishers, 798–816.
- Wilks, Yorick; Webb, Nick; Setzer, Andrea; Hepp, Mark; Catizone, Roberta 2005. Machine learning approaches to human dialogue modelling. – C. J. van Kuppevelt, Laila Dybkjær, Niels Ole Bernsen (Eds.). Advances in Natural Multimodal Dialogue Systems. Text, Speech and Language Technology, 30. Dordrecht: Springer, 355–370. [http://dx.doi.org/10.1007/1-4020-3933-6\\_16](http://dx.doi.org/10.1007/1-4020-3933-6_16)
- Williams, Jason D.; Young, Steve 2007. Partially observable Markov decision processes for spoken dialog systems. – Computer Speech and Language, 21 (2), 393–422. <http://dx.doi.org/10.1016/j.csl.2006.06.008>

**Mare Koit** (Tartu Ülikool) on uurinud dialoogi modelleerimist arvutil.  
[mare.koit@ut.ee](mailto:mare.koit@ut.ee)

# **MODELLING OF CONVERSATIONAL AGENTS IN ARGUMENTATION: CONVERSATION AS UPDATING OF INFORMATION STATES**

**Mare Koit**

University of Tartu

This article considers conversations where one participant is trying to lead his/her partner to a decision: to do an action. The concentration is on the representation of information states of the conversational agent and their update rules. In conversation, the agent moves from one information state to another, applying the update rules. An information state includes a partner model which consists of evaluations which the agent believes his/her partner has assigned to different aspects of the action under consideration. The partner model changes during conversation, according to arguments and counter-arguments presented by the participants. As a practical implementation of the model the present study has in view a dialogue system which is called a communication trainer. Estonian human-human dialogues are considered and the applicability of the model is discussed.

**Keywords:** argumentation dialogue, dialogue model, reasoning model, conversational agent, Estonian