University of Tartu

Department of Philosophy

Making Inferences in case of Attitudes

MA Thesis

Anto Unt

Scientific supervisor Prof. Rein Vihalemm

Tartu 2010

Contents

Introduction

In the Cambridge Dictionary of Philosophy [ed. Robert Audi, Cambridge University Press 1995] the Standard Deontic Logic (SDL) is characterized in the following way:

> The overwhelming verdict among deontic logicians is that SDL genuinely succumbs to the deontic paradoxes. But it is controversial what other approach is best followed to resolve these puzzles. [James E Tomberline, "deontic paradoxes", p 191 in the referred dictionary.]

"Deontic paradoxes" are outrageously absurd consequences following from purported rules that allow deducing new obligations from given ones. The three papers that constitute the current thesis share the idea that the proper solution to this problem is to give up the very idea of deducing new obligations relying on some underlying logic. This idea may be stated in more detail by the following claims that are common to all the three papers.

1) When talking about making inferences in case of obligations, it makes sense to start with the relation "Obligation $O(p)$ is fulfilled if and only if p". (The F-schema.)

2) If there is Obligation(p) and p implies q, the F-schema allows to conclude that not-q implies violation of Obligation(p). It is this conclusion that should be considered as a truly logical one, not the existence of Obligation(q); in fact, there is neither need nor ground to deduce the existence of Obligation(q).

3) The F-schema also allows to conclude that whenever there is Obligation(p) and q implying p, q entails fulfillment of the Obligation(p); in particular, if I am obliged to this or that then I can fulfill the obligation by making either of the disjunct true. It is a strange feature of the standard deontic logic that it does not pay any attention to such kind of reasoning, although it is very common in everyday practice, legal matters included.

4) Propositional attitudes – beliefs and desires – behave in the same way as obligations when it comes to making inferences. To include all of them we have just to use a bit more general formulation, something like, "The attitude $A(p)$ finds positive solution if and only if p"; the latter may be labelled as a "P-schema".

When defending my M.A. Thesis I would like to concentrate on the 3rd, unpublished paper,

hoping to still improve it. Already in its current state the 3rd paper is distinguished mainly by its conscious attempt to enter into a dialog with potential readers. In order to achieve this aim, the following new moments of the 3rd paper may help, I hope.

1) The F-schema is now compared, in detail, to Tarski's T-schema that allows, on one hand, to recall familiar associations and, on the other hand, to demonstrate specific character of attitudes as compared to statements; it will be done in Sec.2 of the paper.
2) A considerable attention will be payed to cases where, contrary to the main claim of the current thesis, deducing new obligations from existing ones seems to take place. In Sec. 3, apparent inferences will be explained as application of certain pragmatic rules that are distinguished from logical rules by their *ad hoc* correctibility – whenever the rules yield unexpected results the applicability of the rules is turned down. The fact that an obligation concerning a particular situation seems to exist whenever the situation qualifies as a relevant instance of a general obligation still seems to require an additional analysis; such an analysis will be presented in Sec. 9.
3) McGee's purported counterexample to *modus ponens* has been analyzed as a suitable example to demonstrate the benefits of the suggested approach.

In addition, some other problems (definining knowledge in a way that conforms with philosophers' intuition about Gettier's examples, counting for opacity of belief contexts) will be briefly discussed. Although these problems might be interesting by themselves, I have found it difficult to discuss them without relying on P-schema, so the discussion will be included in the current paper.

Hoiakuloogika

Ilmunud *Akadeemia,* **1997**:2, 344-357

1. Sissejuhatus: probleem disjunktiivse lubamisega.

Pikka aega olen olnud arvamisel, nagu vist nii mõnedki, et loogika kui selline on ära sõnastatud juba Aristotelese aegadel. Et Uusajal on muidugi formalismi kõvasti edasi arendatud, mis on aga rohkem selleks hea, et ka rumalamad inimesed keerulisemaid arutlusi suudaksid jälgida. See tähendab, et kui võtta mõni Platoni mõttekäik (näiteks: "Kui Sokrates ei usu, et surres läheb ta Jumalate juurde, siis toimib ta valesti mitte surma vastu olles. Kui Sokrates seda usub, siis ei toimi ta valesti mitte surma vastu olles. Seega ta kas usub, et surres läheb ta Jumalate juurde, ja ei toimi valesti mitte surma vastu olles või ta ei usu seda ja toimib valesti mitte surma vastu olles."), siis targematele inimestele on niigi selge, et tegemist on õige järeldusskeemiga (nad saavad aru, et sõna "seega" kasutamine on õigustatud), aga samale arusaamisele võivad jõuda ka vähem targad (näiteks mina), kui nad need laused formaliseerituna kirja panevad ning lausearvutuse reegleid kasutavad.[1] Muidugi ei ole loogika areng, eriti viimasel sajandil, piirdunud lausearvutusega; aga minu aasta-paari tagune üllatus on seotud just nimelt avastusega, et teatud juhtudel ei saa ma aru ka kõige igapäevasemast lausearvutusest – sellest, milliste reeglite järgi mõnikord teisenevad laused, mis sisaldavad sõnu 'või', 'ja' ning 'ei'.[2] Kõige ilmekamaks näiteks on siin disjunktiivne lubamine (loa, mitte lubaduse andmise mõttes), mille üle on viimastel aastakümnetel ka trükisõnas arutletud[3]: kui mulle öeldakse, et ma võin avada ukse või akna, ja ma avan ukse, siis olen ma kindel, et olen toiminud nii nagu lubatud. Kui me lubamisoperaatori tähistame P-ga (sõnast 'permission' või 'permissio' – kumba keelt keegi eelistab), siis väljendab seda arusaamist implikatsioon

(1) $\quad P(a \vee b) \rightarrow P(a)$.

---

[1] Toodud näite korral piisab, kui veenduda, et kehtib
$(\sim U \rightarrow T) \,\&\, (U \rightarrow \sim T) \equiv (U \,\&\, \sim T) \vee (\sim U \,\&\, T)$, kus *U* tähistab väidet "Sokrates usub, et surres läheb ta jumalate juurde" ja *T* väidet "Sokrates toimib valesti mitte surma vastu olles". (Näide pärineb Tõnu Tamme loogikakursusest, kuhu see on võetud dialoogist "Phaidon".)

[2] Seda avastust ei teinud ma iseseisvalt, vaid Madis Kõivu suunamisel; avaldan talle siinkohal tänu mitte üksnes esialgse tõuke, vaid ka pärastiste mõjutuste eest, mis on tähendanud nii innustavaid kui ka kainestavaid märkusi – mõlemad on osutunud tarvilikuks (ja võivad osutuda tarvilikuks ka edaspidi).

[3] Lubamise, käskimise ja keelamisega seotud valdkonda nimetatakse sel puhul deontiliseks loogikaks, mille varasemast ajaloost annab hea ülevaate [Føllesdal, Hilpinen 1970]. Disjunktiivse lubamisega seotud paradoksi kohta võib lugeda artiklist [Hilpinen 1982].

Esimeseks küsimuseks võiks nüüd olla niisuguse implikatsiooni päritolu. Sest disjunktiivsete *väidete* jaoks kehtib ju vastassuunaline implikatsioon ($a \rightarrow a \vee b$) ja ka 'lubamisele' mõneti sarnase võimalikkuse operaatori $\Diamond$ jaoks kehtib traditsioonilises modaalses loogikas just $\Diamond a \rightarrow \Diamond(a \vee b)$, mitte aga (1) analoog. Kuid mitte see pole peamine mure, sest võiks ju ka lihtsalt postuleerida, et lubamise operaatori korral aksioomi (1) lihtsalt kehtib ja kõik. Ebameeldivused tabavad meid aga siis, kui hakkame vaatlema niisugusest aksioomist tulenevaid järeldusi, eeldades lisaks, et lubamise operaatori all kehtib tavaline lausearvutus, mida väljendaks valem:

(2) $(a \equiv b) \rightarrow (P(a) \equiv P(b))$.

Järeldusskeem (2) võib tunduda päris loomulik ja traditsiooniline deontiline loogika sellist reeglit enamasti ka tunnistab[4]; ning see ei sünnita mingeid erilisi probleeme, niikaua kui implikatsiooni (1) *mitte* tunnistada. Kui seda aga teha, siis ei takista meid miski seda rakendamast ka järgmisele aksioomiga (2) seadustatud ekvivalentsusele

(3) $P(a) \equiv P((a \& b) \vee (a \& \sim b))$,

mis annab tulemuseks

(4) $P(a) \rightarrow P(a \& b)$,

ja selle peale on küll põhjust ära kohkuda: tähendab ju (4), et kui mingigi asi on lubatud, siis koos temaga on alati lubatud ka *mistahes* muu asi.

Jätan igaühe enda otsustada, kui süütu *a* ja kui jubeda *b* abil ta soovib implikatsiooni (4) endale ette kujutada. Ja loodan igaühe nõustumist, kui väidan, et (4) ei saa mingil juhul õige järeldusskeem olla.[5] (Vastasel korral võiks karta, et mitte keegi ei söandaks kunagi kellelegi millekski luba anda.) Kuna aga (4) on vahetu järeldus (1) ja (2) kooskehtimisest, siis peaksime me (4) vältimiseks kas (1) või (2) vääraks tunnistama – muidugi kui me ei võta omaks seisukohta, et tavalises keeles ei kehtigi mingit loogikat, nii et siin lihtsalt ei ole mõistlik järjekindlusega üle pakkuda. Viimane võimalus võib küll ülearu kapitulantlik tunduda, kuid paraku ähvardavad (1) või (2) hülgamine samasuguse tulemusega: kui ei kehti (1), millest inimesed alailma lähtuvad, siis tähendab see ju seda, et harilikus keeles pole mingit loogikat. Teisest küljest tundub ka (3) olevat igati mõistlik järeldusskeem: kui lubatud tegude seas on *a*, siis peab ju lubatud tegude seas olema ka kas '*a* ja *b*' või '*a* ja mitte-*b*'.[6]

---

[4] Oma esimeses deontikaalases artiklis lülitas seose (2) oma aksioomide hulka ka von Wright [von Wright 1951]; ta nimetades seda ekstensionaalsuse reegliks.

[5] Olen kohanud ka teistsugust arvamust: (4) võiks väljendada lihtsalt asjaolu, et ainult ühte asja korraga ei saagi lubada, sest mistahes asja lubades lubatakse alati ka lubatu tagajärgi. Paraku ei osuta selles implikatsioonis miski sellele, et *b* peaks mingil viisil *a*-ga seotud olema ('tagajärje' käsitlemine loogiliste vahenditega on üldse problemaatiline); nagu juba rõhutatud, võib *b* olla *mistahes* asi.

[6] Iseäranis peaks olema selge, et tegemist ei ole mingi hägusloogika probleemiga, kus *b* ja mitte-*b* vahel veel mingi ebamäärase ulatusega määramatus võimutseb. Vaadeldav probleem ilmneb juba

Pärast mõningat järelemõtlemist tuleb tunnistada, et loomuliku keele loogilisel analüüsil tuleb sedalaadi ummikuid päris tihti ette; isegi nii tihti, et on alust mõningaid väljapääsuvõimalusi standardseks nimetada – näiteks ideed sel puhul keelest varjatud kahemõttelisust otsida. Niisiis võiks küsida: milles seisneb antud juhul kahemõttelisus? Ja miks need kaks mõtet omavahel kokku ei sobi?

Reegli (2) analoog kehtib traditsioonilises modaalses loogikas ja seega võime öelda, et mingis mõttes on selle tähendus meile tuttav. Seevastu, nagu juba mainitud, on implikatsiooni (1) päritolu mõnevõrra ähmane. Meie eesmärk võiks siis olla seda päritolu täpsustada, lootes selle käigus saada ühtlasi aimu sellest, miks (1) ja (2) omavahel kokku ei sobi. Aga enne kui asuda implikatsiooni (1) põhjendama, küsigem, milline on üldse see valdkond, kus sedalaadi järeldusmalle pruugitakse.

2. Valdkond, kus kasutatakse loogikat, milles võiks kehtida implikatsioon (1).

Väljendiga "*Sinu* loogika järgi tuleb välja, et …" avaldatakse enamasti kahtlust mingi arutelu pädevuse kohta. Siin ilmneb arusaam, et loogika on ebaisikuline ja kui keegi mingeid isiklikke järeldusskeeme juurutab, siis küllap ta eksib. Analoogilist umbusku võiks sünnitada väide, et 'lubamise' korral tähendab sõna 'või' midagi teistsugust võrreldes oma tavalise tähendusega – seisneb ju loogiliste konstantide mõte just nende universaalsuses, nende **mitte**-sõltumises konkreetsest kasutamisjuhust. Seepärast küsigem alustuseks, kas seost (1) sisaldava loogika kasutusala ei võiks laiem olla.

Kui lubamises on nähtud võimalikkuse analoogi, siis *niisuguse* käsitluse korral on paratamatuse analoogiks käskimine (või kohustamine).[7] Kui ka kohustuse korral disjunktsiooni vastu huvi tunda, siis *nii* värvikat paradoksi kui lubamise korral me ei leia; aga midagi häirivat on ka siin. Nimelt, kui mulle öeldakse, et ma olen kohustatud tegema *a* või *b* (sooritama inglise või saksa keele eksami), siis olen ma kindel, et olen oma kohustuse täitnud, kui ma teen ükskõik kumma teo, kas ainult *a* või ainult *b*. Küsimus on nüüd selles, kuidas see asjaolu peaks kajastuma deontilises loogikas. Sest *modaalsest* loogikast me midagi analoogilist ei leia: kui on teada, et 'paratamatult *a* või *b*', siis *a* ega *b* kohta eraldi võetuna ei järeldu siit midagi.

Disjunktiivse kohustusega on tegeldud ja üht lahendust on nähtud võimaluses tõlgendada lauset 'kohustus *p*' kui ähvardust 'mitte-*p* implitseerib karistuse'. Ja tõepoolest, disjunktiivse kohustuse tähendus on sel kombel tabatud, ehkki nüüd kipub kaotsi minema kohustuse seos lubamisega. Seetõttu

---

tavalises kahevalentses loogikas, millega me oma käsitluses ka piirdume.

[7] 'Paratamatu, et *p*' on ekvivalentne väljendiga 'ei ole võimalik, et mitte-*p*'. Analoogiliselt, 'kohustuslik, et *p*' on ekvivalentne väljendiga 'ei ole lubatud, et mitte-*p*'.

vajab selline lähenemine edasisi täpsustusi, mille käigus aga säilib ‘karistuse’ või mõne muu just deontikasse kuuluva mõiste eriline roll.

Sellega oleme jõudnud käesoleva kirjutise esimese ideeni: miks piirduda deontikaga, kui disjunktiivsele kohustusele analoogiline nähtus leiab aset ka disjunktiivsete uskumuste ja soovimiste korral? Sest öeldes ‘ma usun, et *a* või *b*’ ütlen ma seeläbi, et pean võimalikuks nii *a*-d kui ka *b*-d; ja öeldes ‘ma tahan *a*-d või *b*-d’ ütlen ma seeläbi, et aktsepteerin oma soovide rahuldamisena nii *a*-d kui ka *b*-d.[8] Niisiis katsume järgnevalt vaadelda neid kolme koos, nimetades neid ühise nimega ‘hoiak’ ja tähistades vastavat operaatorit tähega A (sõnast ‘attitude’). Nagu nägime disjunktiivse hoiaku analüüsil, on igal hoiakul oma paariline[9]; esitame hoiakud koos oma paarilisega järgmise tabelina:

Hoiak *Ap*	Hoiaku *Ap* paariline *A′p*	Tõeväärtuslik hinnang p-ga seotud hoiakule (T_p)
Usun, et *p*.	Pean võimalikuks, et *p*.	*p*-ga seotud uskumus vastab tõele.
Soovin, et *p*.	Aktsepteerin soovitavana, et *p*.	*p*-ga seotud soov on rahuldatud.
Kohustus, et *p*.	Kohustuse *p* täidetus.	*p*-ga seotud kohustus on täidetud.

Siia tabelile on lisatud veel kolmas veerg, kus on kirjas tõeväärtuslik hinnang hoiakule, mida tähistab T_p. Just selle täiendava tõeväärtuse abil loodame määratleda hoiakuloogikat niimoodi, et osutub võimalikuks selgitada (muuhulgas ka) disjunktiivse hoiaku tähendust.

3. Hoiakute semantiline interpretatsioon.

Mingi spetsiifilise loogika esitamise kõige autoriteetsemaks viisiks on alustada sobivate aksioomide

[8] Analüütilises filosoofias on kombeks ‘uskumust’ [belief] ja ‘tahtmist’ [desire] hõlmata terminiga ‘propositsionaalne hoiak’ [propositional attitude]. (Siit pärineb ka kirjutise pealkiri.) Väljendit ‘propositsionaalne hoiak’ kasutavad tihti ka need, kes ei ole nõus käsitlema neid hoiakuid kui relatsiooni subjekti ja propositsiooni vahel. Vastuseisu propositsioonile põhjustab asjaolu, et ‘propositsiooni’ peaksid iseloomustama kindlad tõetingimused, uskumuste korral on aga tegemist nn. läbipaistmatu kontekstiga, kus üldjuhul ei saa samaväärseid mõisteid üksteisega asendada. (Sest see, kelle uskumustest me räägime ei pruugi nende mõistete samaväärsusest teadlik olla.) Selles mõttes erinevad uskumused kohustustest: inimesel võivad olla küll erinevad uskumused Cicero ja Tulluse kohta, kui ta aga on kohustatud Tullusele teate edasi andma, siis ei ole kombeks eristada seda kohustusest edastada teade Cicerole. (‘Soovimine’ näib selles suhtes käituvat vahepealsel viisil.) Kui ma käesolevas kirjutises üritan ‘uskumusi’ ja ‘kohustusi’ ühisel viisil käsitleda, siis eeldan, et sedalaadi probleeme saab lahendada lahus meid siin huvitavast lausearvutusest.

[9] Kohmakavõitu väljend ‘paariline’ on tingitud asjaolust, et siin ei saa kasutada konkreetsema tähendusega sõna ‘duaalne’; viimase kasutamine peaks tähendama seda, et paarilisi saab ühesuguse eeskirja abil teineteise kaudu esitatada (nagu näiteks ‘paratamatust’ ja ‘võimalikkust’). Kui meie järgnev käsitlus peaks õige olema, siis siin vaadeldatavte *A* ja *A′* vahel niisugune sümmeetriline suhe puudub.

väljapakkumisest. Niisuguse lähenemisviisi eeliseks on see, et ei ole tarvis siduda end mingi konkreetse interpretatsiooniga, mistõttu on võimalik saavutatada käsitluse suurem abstraktsus. Sel juhul keskendub loogika tuletusreeglitele ja niisugust käsitlust nimetatakse süntaktiliseks. Süntaktiline lähenemisviis võimaldab meil hoiduda tülikatest metafüüsilistest küsimustest nagu: mis laadi asjade olemasolu me tunnistame?

Kui meil on aga tegemist kahemõttelise olukorraga (nagu me kahtlustame), siis ei pruugi sobivate aksioomide väljavalimine olla sugugi kerge ülesanne: me ei pruugi teada, kust kahe mõtte vaheline piir läheb. Seepärast ma loobun sobivate aksioomide äraarvamisest (ehkki olen seda omaette üritanud) ja asun kohe semantilise interpretatsiooni juurde, mis loogikas ei tähenda midagi muud kui opereerimist tõeväärtustega. Üheks tõeväärtuse kandjaks on seejuures propositsioon *p*, millele hoiak *Ap* on suunatud.[10] (Et tegemist on lihtsustusega, mille rakendatavuspiirid on ebamäärased, sellest oli juttu märkuses 8.) Aga lisaks on tarvis veel teist sorti tõeväärtust, milles kajastuks antud hoiaku spetsiifika, ja selleks ongi tõeväärtuslik hinnang hoiakule, nii nagu tabelis toodud.

Pärast neid selgitusi paneme kirja järgmised definitsioonivalemid (tegemist on selle kirjutise teise ja ka viimase ideega – ülejäänu on vaid selgitused siia juurde):

(D1) $\diamondsuit_p p =_{def} p \rightarrow T_p$

(D2) $\boxdot p =_{def} \sim p \rightarrow \sim T$

(D3) $A'p =_{def} \diamondsuit_p p$

(D4) $Ap =_{def} \diamondsuit_p p \,\&\, \boxdot p$

(D4') $Ap =_{def} p \equiv T_p$

Neid valemeid kommenteerima asudes tuleks kõigepealt põhjendada indeksi *p* kasutamist. Tõtt öelda meeldiksid need valemid mulle palju rohkem, kui siin õnnestuks ilma indeksiteta hakkama saada, aga kahjuks pole see võimalik. Vaatame definitsioone (D1) ja (D3). Võttes hoiakuks uskumise, võiks neid koos lugeda nii: kui ma pean võimalikuks, et *p*, ja kui *p* osutub tõeseks, siis minu *p-ga seotud* uskumus on tõene. Just '*p*-ga seotud uskumus' ja mitte 'minu uskumustesüsteem tervikuna' (viimast tähistaks ilma indeksita T). Küll võib aga indeksi ära jätta siis, kui tegemist on *mitte*-T_p-ga – ühe uskumuse vääraks tunnistamine tähendab terve uskumussüsteemi vääraks tunnistamist. (Viimane väide tähendab uskumuste süsteemi tõlgendamist üksikuskumuste konjunktsioonina. Võimalus vastuoluliste uskumuste koosesinemiseks jääb sel juhul muidugi vaatluse alt välja. Tegemist on taas sedalaadi lihtsustusega, millest oli juttu märkuses 8.) Deontilise loogika korral, kus mitte-T-ks on

[10] Siinkohal oleks täpsem eristada *p* kaht tüüpi tõväärtust, olenevalt sellest, kas hoiak on pidev või perfektne (ajavormi, mitte täiuslikkuse mõttes). Kui tähistada t_p-ga ajahetke, mil hinnatakse propositsiooni tõesust, ja t_a-ga ajahetke, mil hinnatakse hoiaku tõesust, siis oleks 'p' täpsem lugemine järgmine: perfektse hoiaku korral '$\exists t_p$ ($t_p<t_a$) & p' ning pideva hoiaku korral '$\forall t_p$ ($t_p=t_a$) → p'. Kuivõrd ma ei näe ära, et see vahetegemine mõjutaks käesolevas kirjutises vaadeldavate probleemide

kohustuste mittetäitmine, võib selle asemel rääkida ka 'karistamisest' (just nii nagu seda tehaksegi) ja pääseda sel kombel tülikast indekseerimisest. Kuid ma kahtlen, kas ainuüksi 'karistamise' abil on võimalik deontika täielikku semantikat esitada.

Asume nüüd oma valemite selgitamise juurde ja alustame definitsioonist (D4), mis võib esmapilgul liigagi triviaalne näida. Ütleb see ju lihtsalt järgmist (eksemplariseerides hoiakut tahtmisena): 'ma tahan, et p' tähendab definitsiooni kohaselt 'kui p, siis on mu (see) tahtmine täidetud, ja kui mitte-p, siis mu tahtmised ei ole täidetud'. Siiski loeksin ma selle definitsiooni endastmõistetavust pigem tema eeliseks: kui see definitsioon oleks komplitseeritum, siis võiks ju igaüks vastu vaielda, et tema oma hoiakuid küll nii keerulisel viisil ei määratle. Aga mis siis ikkagi on selle definitsiooni (taga)mõte? See ilmneb kohe, kui vaadelda hoiakut, mis on suunatud disjunktiivsele propositsioonile $p=q\vee r$; lähtudes definitsioonivalemitest ja kasutades üksnes tavalist lausearvutust, leiame sel puhul:

(5) $A\,(\,q\vee r\,) \equiv \diamond_{\mathrm{p}} q \;\&\; \diamond_{\mathrm{p}} r \;\&\; \boxdot(q\vee r)$.

Ja siin sisaldub täpselt see, millest oli juttu punktis 2 seoses disjunktiivse hoiaku tähendusega: 'ma usun, et *q* või *r*' implitseerib 'ma pean võimalikuks, et q' ja 'ma pean võimalikuks, et r'. Aga lisaks punktis 2 juba öeldule on valemist (5) näha ka see, milline väide tuleb neile kahele 'võimalikuks pidamisele' lisada selleks, et implikatsiooni asemel saaks rääkida ekvivalentsist; selleks väiteks on $\boxdot(q\vee r)$, mida võiks lugeda näiteks nii: 'kui ei ole tõsi, et *q* või *r*, siis mu usk, et *q* või *r* on väär'. Kui me aga prooviksime disjunktiivse hoiaku tähendust puhtintuitiivselt selgitada, siis võib kergesti jõuda umbes niisuguse väiteni: " 'Ma usun, et *q* või *r*' tähendab seda, et 'ma pean võimalikuks *q*-d' ja 'ma pean võimalikuks *r*-i' ja … 'ma usun, et *q* või *r*' ". See, mille poolest see viimane 'uskumine' erineb esimesest 'uskumisest' saab aga selgeks definitsioonist (D4) – esimese uskumise puhul on tegemist loogilise ekvivalentsi ehk bikonditsionaaliga, teisel juhul ühe seda ekvivalentsi moodustava konditsionaali ehk implikatsiooniga.

Soov seletada disjunktiivse hoiaku tähendust on olnud senise käsitluse peamiseks eesmärgiks. Aga muidugi tuleb kontrollida ka seda, ega leitud seletus muudel puhkudel mingeid liiga absurdseid tulemusi ei anna. Vaatame siis kõigepealt hoiakut, mis on suunatud konjunktsioonile $p=q\&r$:

(6) $A\,(\,q\&r\,) \equiv \boxdot q \;\&\; \boxdot r \;\&\; \diamond_{\mathrm{p}}(q\&r)$.

Valem (6) pole küll ehk nii sisukas kui valem (5), aga ei saa ka öelda, et siin midagi valesti oleks. Huvitavama tulemuse saame aga siis, kui vaatleme hoiakut, mis on suunatud implikatsioonile $p = q\rightarrow r \equiv$

---

käsitlust, siis rohkem ma sellel ei peatu.

$\sim q \vee r$:

(7) $A\,(\,q \rightarrow r\,) \equiv \diamondsuit_{\mathrm{p}} \sim q \;\&\; \diamondsuit_{\mathrm{p}} r \;\&\; \boxdot (q \rightarrow r)$.

Ma nimelt väidaksin, et valem (7) lähendab implikatsiooni sammukese võrra tavalisele 'kui … siis' lausele: öeldes '(ma usun, et) kui *q*, siis *r*' ütlen ma vastavalt valemile (7) ka seda, et 'ma pean võimalikuks, et mitte-*q*' ja 'ma pean võimalikuks, et *r*'. See, et tavaline 'kui … siis' lause sisaldab endas ka tingimust $\diamondsuit_{\mathrm{p}} \sim q$ on minu meelest eriti ilmne eesti keeles, kus sõna 'kui' võib samas positsioonis väljendada nii tingimuslikkust kui ka ajamäärust. Ja tingimuslikkust väljendab ta üksnes siis, kui on täidetud $\diamondsuit_{\mathrm{p}} \sim q$ (ma pean võimalikuks, et mitte-*q*). Selles veendumiseks võrrelge näiteks järgmisi lauseid: "kui homme hommikul päike tõuseb, siis läheme kalale" ja "kui homme hommikul päike paistab, siis läheme kalale".

Kuidas on aga lood eitusega? Moodustatakse ju loomulikus keeles 'uskumise' ja 'tahtmise' korral eitus mõnevõrra iseäralikul viisil. Võtame näiteks 'uskumise'. Väljendid 'ma usun, et p' ja 'mul on uskumus, et p' näivad olevat samaväärsed. Samal ajal lause 'ma ei usu, et *p*' ei tähenda sedasama, mis 'mul puudub uskumus, et *p*', esimene on pigem ekvivalentne lausega 'ma usun, et mitte-*p*'. Analoogiliselt eelnevaga, 'ma ei taha, et *p*' tähendab pigem 'ma tahan, et mitte-*p*'. ('Kohustamise' tarvis eesti keeles analoogilisel viisil käituv sõna puudub, küll aga sobivad siia ritta inglise keele modaalsed verbid 'must', 'should' ja 'ought to'.) Kas meie definitsioonid kajastavad seda asjaolu? Vahetu lausearvutus annab:

(8) $\sim Ap = \sim p \equiv \mathrm{T_p}$

ja

(9) $A{\sim}p = \sim p \equiv \mathrm{T_{\sim p}}$

Seega võime väita, et *kui* kehtib seos

(10) $\mathrm{T_p} = \mathrm{T_{\sim p}}$,

siis kehtib hoiaku eituse jaoks valem

(11) $\sim Ap \equiv A{\sim}p$,

mis täpselt 'ei usu' ja 'ei taha' tähendust edasi annab.

Mida tähendab seos (10)? Ei midagi rohkemat kui seda, et mitte-*p*-ga seotud hoiak on määratud samadel

tingimustel kui *p*-ga seotud hoiak. Kui nüüd võrrelda eituse moodustamist sõnapaari {'uskuma','tahtma'} korral üheltpoolt ja sõnapaari {'arvama','soovima'} korral teiseltpoolt, siis jääb mulje, et esimene sõnapaar tunnistab seost (10) tingimusteta (sest kehtib (11)), teine sõnapaar jätab selle küsimuse aga lahtiseks: öeldes 'ma ei arva, et *p*' ei pea ma samal ajal arvama, et mitte-*p*, mul võib lihtsalt *p* kohta üldse igasugune arvamus puududa. Analoogiliselt 'ma ei soovi' väljendab pigem soovi puudumist kui vastupidist soovi. Millisele loogikale allub siis sõnapaar 'arvama','soovima'? Kahtlustan, et üks võimalik vastus oleks: silmakirjalikule loogikale. Et seda kahtlustust põhjendada, tuleb hoiakuloogikat võrrelda tavalise modaalloogikaga.

4. Hoiakuloogika võrdlus modaalloogikaga, unustamata deontikat.

Rääkides hoiakutega seoses 'semantilisest interpretatsioonist', laenasin ma selle mõiste modaalloogikast. (Vaata näiteks [Cresswell, Hughes 1996].) Modaalloogika semantiline interpretatsioon on kurikuulus selle poolest, et enamus asjaosalisi on nõus kahe väitega, mille rahulik kooseksisteerimine on raskendatud. Nendeks väideteks on esiteks, et semantiline interpretatsioon teeb paljud modaalloogika probleemid selgemaks ja arusaadavamaks, ning teiseks, et see interpretatsioon rajaneb mõistel, mis ise ei ole sugugi selge ega arusaadav. Tolleks mõisteks on 'võimalik maailm' ja lause 'võimalik, et *p*' semantiliseks interpretatsiooniks on 'leidub võimalik maailm, kus *p*'. Kusjuures selline tõlgendus *ei ole* tsirkulaarne (mõiste määratlemine tema enda kaudu), kui võtta võimalikku maailma kui analüüsimatut primitiivi. Tõsi, vaid vähesed inimesed on rahul olukorraga, kus mõiste 'võimalik maailm' edasisele analüüsile ei allu, ülejäänud on sunnitud 'võimalike maailmade' teemalisse väitlusse astuma. Aga just see, et paistab nagu oleksid nad selleks sunnitud, näitab, et 'võimalike maailmade realism' *on* modaalloogika jaoks (teadaolevatest) kõige loomulikum tõlgendusviis. Antud juhul pole minu huvi siiski mitte kaitsta 'võimalikke maailmu', vaid üksnes tagasihoidlikumat väidet, et modaalloogika suunab meid 'võimalikkuste reaalsuse' tunnistamisele. Sest see näib olevat modaalloogika üheks eesmärgiks – käsitleda olukorda, kus võimalikkused on päriselt olemas. (Loogika puhtsüntaktiline esitusviis võimaldab muidugi alati eesmärke varjata.)

Just selles suhtes vastandubki modaalloogikale hoiakuloogika: kalduvuse asemel tunnistada võimalikkuste iseseisvat eksistentsi leiame siit valmisoleku anda hinnang ettetulevatele olukordadele. Modaalloogika seisukohalt on olemas võimalikkused ja me kogeme ühte neist. Hoiakuloogikas kogeme me kõigepealt midagi ja seejärel anname sellele hinnangu; ning hinnanguks võib olla ka 'võimatu' – hoiakuloogikas tähendab see vaid mõne uskumuse vääraks osutumist. Samasuguse erinevuse võib sõnastada ka 'tahtmiste' korral: hoiakuloogika seisukohalt me mitte ei vali välja üht olukorda aktsepteeritavate olukordade hulgast, vaid oleme valmis igas ettetulevas olukorras hindama, kas me seda aktsepteeritme. Modaalloogikat võiks siis nimetadaektensionaalseks ja hoiakuloogikat intensionaalseks.

Kui nüüd aga taas meenutada sõnapaari {‘arvan’,‘soovin’}, siis näeme, et nende sõnade käitumises ilmneb mõlemale loogikale omaseid jooni – disjunktsiooni puhul käituvad nad nii, nagu oleks tegemist hoiakuga, eituse puhul (võivad käituda ka) nii, nagu oleks tegemist ekstensionaalse loogikaga (st. nagu oleks soovidel ja arvamustel subjektist sõltumatu eksistents). Kui ma neid sõnu silmakirjalikeks klassifitseerin, siis lähtun oletusest, et nende õige koht on hoiakute hulgas, kus nad ei erineks vastavalt tahtmisest ja uksumisest. Kui aga juhtub, et ma soovin (tahan?) jätta silmakirjalikku muljet, nagu poleks mul omaenda soovide ja arvamustega mingit pistmist, siis on seda sobiv teha sõnade abil, mille eitus moodustatakse ekstensionaalse loogika reeglite järgi.

Kuhu kuulub aga deontiline loogika? Siiamaani proovisin ‘kohustust’ käsitleda sama malli järgi kui ‘uskumist’ ja ‘tahtmist’; nüüd võiksin seda põhjendada sellega, et inimesel endal peab olema võimalus hinnata, kas ta on käitunud lubatud viisil, mis määrab ära deontika kuulumise hoiakuloogikasse. Paraku tuleb tunnistada, et ‘kohustuste’ taotlus kuuluda ekstensionaalse loogika sfääri on samuti õigustatud: ma ei tahaks päriselt eitada ka ‘kohustuste’ subjektivälist eksistentsi. Kui nii, siis on põhjendatud ‘kohustuste’ ja ‘lubamiste’ käitumine kahtede erinevate reeglite järgi, mispuhul paradokside esinemine on loomulik. Kahe loogika olemasolu nähes me vaid oskame paradokse paremini analüüsida. Näide, mida analüüsida, on meil juba olemas: see on esimeses punktis saadud paradoksaalne tulemus (4). Niisiis vaatame nüüd, kuhu kuuluvad implikatsioonini (4) viinud eeldused. Ja me näeme kohe, et seos (1) kuulub hoiakuloogikasse, seos (2) aga ekstensionaalsesse loogikasse. Seos (1) võimaldab küsimust “kas ma käitun lubatuga kooskõlaliselt?”, seos (2) aga küsimust “missugused lubatavad olukorrad on olemas?”. Kumbagi küsimisviisi ei ole mõtet ära keelata, alusetu on vaid nende läbisegi kasutamine. Ja kui loogikaid eraldi kasutada, siis ei teki ka mingeid paradokse: hoiakuloogikas ei kehti (2), sest puudub võimalus jõuda $\diamondsuit_a$ juurest $\diamondsuit_{a\&b}$ juurde; ja üheski ekstensionaalses loogikas ei kehti niisugust seost nagu (1).

Kirjandus.

Cresswell, M.J., Hughes, G.E. 1996. A New Introduction to Modal Logic. London, Routledge.

Føllesdal, Dagfinn, Hilpinen, Risto 1970. Deontic Logic: an Introduction. – in: Deontic Logic: Introductory and Systematic Readings. Reidel Publishing Company, Dordrecht-Holland. P. 1-35.

Hilpinen, Risto 1982. Disjunctive Permissions and Conditionals with Disjunctive Antecedents. Acta Philosophica Fennica Vol.35. P.175-194.

von Wright, Georg Henrik 1957. Deontic Logic. – His *Logical Studies*. London: Routledge & Kegan Paul, pp. 58-74. [Esmakordselt ilmunud 1951]

Making Inferences in case of Attitudes

Not published yet

0. A remark about using the word 'attitude'

The general aim of this paper is to comment on making inferences that concerns both propositional attitudes (like belief and desire) and deontic concepts (like obligation). The word 'attitude' of the title is meant to cover them all. I use it just because I have not been able to figure out anything better, I am not sure it is the best choice one can make.

1. Two patterns of reasoning in case of obligations

Consider two examples,

(1) There is an obligation to tell truth to anybody. Therefore, there is an obligation to tell truth to my little brother.

(2) There is an obligation for students to pass an examination in German or French. If I pass an examination in French, I therefore fulfill this obligation.

In most, if not all, attempts to formalize deontic reasoning, it is the cases of the first type – deducing new orders from given ones – that are assumed to be related to some underlying deontic logic, whereas questions like fulfilling disjunctive obligations are, if mentioned at all, explained as some pragmatic addition to a true logic. In this paper, I will try to reverse this conception, claiming that it is namely (2) that may be regarded as an instance of purely logical reasoning (as an argument that is correct due to a generally valid pattern it follows), whereas the status of (1) will be found to be controversial.

Let us assume, contrary to the above claim, that (1) follows from a general pattern of deontic reasoning, in which case the only pattern I can think of is,

(3) $(p \rightarrow q) \rightarrow (\text{Obligation}(p) \rightarrow \text{Obligation}(q))$.

(It is also a theorem of the standard deontic logic.) The example (2), in turn, suggests that in this case the inference pattern might be

(4) $(p \rightarrow q) \rightarrow (p \rightarrow \text{Obligation}(q)$ is fulfilled).

Now, at first sight one might suspect that together (3) and (4) give rise to a tonk-like conclusion, as (3) allows to deduce a disjunctive obligation from a simple one, like,

(5) I must slip the letter in a mailbox. Therefore, I must slip it in a mailbox or burn it.

and (4) allows to infer

(6) If I burn the letter, I fulfill the obligation "Slip the letter in a mailbox or burn it!".

On a closer look, however, the conclusion we can actually make with (3),(4) is not that absurd, as far as we cannot infer, relying on the rule (4), that the obligation "Slip the letter in a mailbox!" is fulfilled when I burn the letter, it is just the obligation "Slip the letter in a mailbox or burn it" that is, and there is no way to infer the former fulfillment from the latter one.

Still, even without direct absurdity, (5) seems to be a weird way of reasoning, and this fact alone might cast doubt on (3) as a generally valid principle.[1] Besides, there are directly absurd results, too, that depend straightforwardly on (3), e.g. the paradox of the knower.[2]

If one agrees that knowledge that p – K(p) – implies p then due to (3),

(7) Obligation(K(p)) → Obligation(p) .

Now, if 'Smith set warehouse on fire' expresses a true proposition, and Jones, a guard of the warehouse, is obliged to know it if it has happened, then (7) allows to infer that Smith was obliged to set warehouse on fire.

So, there are several reasons to disbelieve (3). At the same time, I have no idea what principle, if not (3), could explain the inference (1) as resulting from pure logic. In such a situation, I suggest to postpone attempts to explain (1) for a while (in fact, until Sec.3) and to switch, instead, to a different kind of inference one can make in the little-brother-case, an inference that uses the same concept that was helpful in the this-or-that-examination case (2):

(8) If I do not tell truth to my little brother, I do not fulfill the obligation 'Tell truth to anybody!'.

In a significant sense, (8) is able to replace (1) – I tell truth to my little brother, if I do, not in order to follow some deduced obligation, but in order not to violate the original one.

Now, it is rather natural to assume that (8) is an instance of

(9) $(p \rightarrow \neg q) \rightarrow (p \rightarrow$ Obligation(q) is not fulfilled) .

Having two inference patterns (4) and (9), it is not hard to see that they are consequences of, correspondingly, 'If q then the Obligation(q) is fulfilled' and 'If not-q then the Obligation(q) is not fulfilled'; or, combining the two,

(10) Obligation(q) is fulfilled if and only if q.

Although we reached (10) by a sequence of guesses as to just one possible basis establishing

[1] Historically, the example (5) was used by Alf Ross (Ross 1941) as an objection to (3). Logicians usually do not worry much about this example. Von Wright, e.g., insists that there is nothing paradoxical here, as far as we should construe the entailment of an obligation as a suggestion to a rational norm-giver not to impose a contrary-to-entailment obligation, i.e. in this case not to prohibit the action described as 'slipping the letter in a mailbox or burning it'. I will return to von Wright's interpretation of entailment in the next section.

[2] Åqvist 1967.

both (2) and (8), it looks plausible in itself and, as such, deserves a closer examination.

2. The F-schema and the T-schema.

The relation (10) – call it F-schema – reminds, of course, Tarski's T-schema, especially when the latter is taken in the form,

(11) Statement(q) is true if and only if q .

Still, a distinctive feature of the T-schema (11), as compared to the F-schema (10), is that in (11), Statement(q) and q are of the same type – both are descriptive or assertoric; it is this feature, I guess, that allows to regard the T-schema as a constraint on theories of truth, or even as *the* theory of truth. In F-schema, on the other hand, there is still a description on the right side, but an obligation to the left of the 'fulfilled', and this is likely the reason why it is hard to imagine (10), by analogy, to be related to some theory of fulfillment.

However, the T-schema has been also used in the opposite direction, as giving the meaning of a statement through its truth conditions[3], and it is this approach that seems to have a useful analogue in case of the F-schema – I can see nothing wrong about the claim that understanding an obligation amounts to knowing its fulfillment conditions.[4]

My primary interest is to excplicate the inferential connections between entities to the left of 'fulfilled' in (10), so I have to rely on certain formal connections for the metalanguage to the right of 'if and only if'. In fact, I assume that customary logic works there, and at least for some candidates to the role of metalanguage this assumpiton is well justified: if the propositions of the metalanguage are taken to be sets of possible worlds then implication $p \rightarrow q$ may be construed as the set of p-worlds forming a subset of the q-worlds, just like conjuction of p and q may be interpreted as the intersections of the sets and disjunction as their union. Besides, the possible-world vocabulary has shown itself as a useful tool for illustrative purposes, so I will adopt it for a while (untill Sec 8), although in most cases I really depend only on inferential relations operating in the metalanguage.

So, the Statement(q) is true in q-worlds and only there. Similarly, the Obligation(q) is fulfilled in q-worlds and only there – the procedure that gives meaning to an obligation reminding the procedure that gives meaning to a statement. An additional advantage, in the latter case, is that we need not worry about a problem with the 'proposition as a set of worlds'-approach that arises in case of statements, namely that all necessary statements will share the same meaning (the set of all possible worlds), just like impossible statements will share their meaning (the empty set of worlds), since obligations – and for that matter, attitudes in general – characteristically make sense only if applied to contingent sentences. (Beliefs may give rise to additional problems, but they will do so anyway; I will come to some of the problems in Secs. 7 and 8.)

Now, there is an important difference when employing the two schemas: while statements, in

[3] Cf, e.g. "Tarski could take translation as syntactically specified, and go on to define truth. But in application to a natural language it makes more sense to assume a partial understanding of truth, and use the theory to throw light on meaning, interpretation, and translation." (Davidson 1977: 204)

[4] Cf, e.g. "[…] the sense of a command is determined by knowing what constitutes obedience to it and what disobedience;" (Dummett 1959: 8).

a sense, come with worlds – one is entitled to make the Statement(q) whenever he is dwelling a q-world –, the existence of the Obligation(q) in a particular world – whatever it might mean – has nothing to do with the q-ness of the world. We may still say, if we like, that for an arbitrary q, the Obligation(q) is fulfilled in a q-world, but then we must give this saying a conditional reading, something like "If Obligation(q) were imposed in a q-world then this obligation would be fulfilled in this world". E.g., we may accept,

(12) 'Obligation(K(p)) is fulfilled' → 'Obligation(p) is fulfilled',

but then the right side of the implication should be read, "If the Obligation(p) were imposed it would be fulfilled"; there is no way to conclude that the Obligation(p) has been actually imposed in a world where Obligation(K(p)) is fulfilled.

It is rather obvious, then, that the existence of new obligations cannot be deduced from a set of existing obligations, the fact that is explicitly admitted, e.g., by von Wright:[5] "That one norm *entails* another norm does *not* mean that if the first "exists", then the second "exists" too". Nevertheless, he wants to make sense of entailment in the spirit of (3), regarding it as a consistency condition imposed on an ideal system of norms: "That an obligation to mail a letter entails an obligation to mail or burn it means "only" that the first is (would be) incompatible with a permission to leave the letter unmailed and do something else."[6] Thus, he defines the deontic entailment in terms of a consistent set of norms, and the latter in terms of a situation where all the norms of a given set are fulfilled. In such a situation, Smith does not put the warehouse on fire and the paradox of the knower does not arise. Similarly, Jones obeys his duty to visit his mother, and Chisholm's paradox (Chisholm 1963) does not arise either.

The analysis on the basis of F-schema (11) conforms to von Wright's argument, in a sense: once we have a valid implication $p \rightarrow q$, together with two obligations Obligation(p) and Obligation(not-q), these obligations are "incompatible" in that they cannot be simultaneously fulfilled.[7] This "incompatibility", however, need not signal about any flaw in the system of obligations, but may just reflect the fact that there are obligations that are triggered by violation of some other norm. So, the fact that a particular obligation is fulfilled (Jones knows that Smith has set the warehouse on fire, or $Jones_2$ informs her mother that he is not coming) entails, indeed, the fact that some other obligation (Smith's obligation not to set warehouse on fire, or $Jones_2$'s obligation to visit her mother) has been already violated, or will be violated. This "incompatibility", however, does not mean much, and neither does the "entailment", so it hardly deserves any attention.

Still I cannot ignore cases where a rather customary sort of entailment seems to take place, like in (1): it is a received view, I guess, that their exists the obligation to tell truth to my little brother as a consequence of there being the obligation to tell truth to anybody.

In the next section I will start undermining this conviction.

---

[5] Von Wright 1993: 45.
[6] *Ibid*: 52.
[7] If Obligation(p) is fulfilled then, according to (11), p; the latter, according to our assumption, entails q; this, again according to (11), entails that Obligation(not-q) is not fulfilled.

3. Introducing new obligations via inferences.

Before coming to differences between statements and obligations, let us start with some similarities concerning their usage in ordinary language. E.g., it is not just (5) that sounds strange, its counterpart in case of statements does not sound better,

(13) He slipped the letter in a mailbox. Therefore, he slipped it in a mailbox or burned it.

although it would be a perfectly normal thing to say that

(14) The letter he wrote cannot be found anywhere. Therefore, he slipped it in a mailbox or burned it.

The reason why (13) sounds weird is, I think, that introducing disjunction requires, in ordinary language, some motivation quite apart from its formal introduction rule, and this motivation is typically the insufficiency of reasons to choose between the disjuncts. Thus, I say 'Wednesday or Thursday' if I know that something happened after Tuesday and before Friday, and I describe a thing as 'green or blue' if I cannot decide which word fits better for a particular color.[8] There are only rare cases when the introduction rule for disjunction is actually used in ordinary language, i.e. when I say 'p or q', although I could have said just 'p': (*a*) when it does not matter, in the given context, whether p or q, or (*b*) it is my deliberate choice not to unveil which one, p or q, I consider to be true.

Now, similar considerations seem to govern not just the case of the introduction rule for disjunction, but a general case of making use of a valid implication $p \rightarrow q$ – one may say q, instead of p, only if (*a*) the expressive power of q, in a given context, is no weaker than that of p, or (*b*) dimming is utterer's deliberate intention. The same restrictions hold when it comes to making utterances about obligations – one can talk about Obligation(q), instead of Obligation(p) (assuming, as previously, $p \rightarrow q$), only if the need to fulfill Obligation(q) is all that matters in the given context. (Deliberate dimming seems to have no analogue in case of obligations.) What we have to assume, then, to explain the apparent universality of examples like (1), is that in such cases, the utterance itself creates the required context. And indeed, it is hard to imagine somebody uttering 'I have to tell truth to my little brother', unless he is in the situation where he has to make up his mind whether to tell truth to his little brother or not. The claim, then, is that one way to explain why 'I have to tell truth to my little brother' seems to be a universal consequence of 'I have to tell truth to anybody' is to suggest that it is reasonable to utter the former only when telling truth to one's little brother is the issue, and if so, then referring to such, strictly speaking non-existing obligation is as good explanation to one's behavior as referring to the original obligation.

I admit that, "But I must tell truth to a particular person, once I must tell truth to any person, and there is no need to think about issues on board to see this" would be a plausible objection to the above explanation, and I will discuss it from a different angle in Sec. 9. Yet there are

[8] Sometimes it is claimed that although a person may utter 'p or q' because he cannot decide between p and q, it is ultimately (at least) one of the disjuncts that makes the utterance true. (Perhaps the person just does not remember which one.) The above examples have been deliberately chosen as counterexamples to this claim: in case of 'green' and 'blue', there need not exist any objective separator between the two color predicates; in case of time, quantum indeterminacy around Wednesday midnight may exclude the possibility to decide between the two dates even in case of a momentary event.

other examples where the current explanation seems to do its job properly. E.g., when I have to pass an examination next Thursday, does it imply that I have an obligation to pass it next Wednesday or Thursday? But what about the obligation to pass it next week? If you agree that the latter is a more plausible conclusion than the former, I see no alternative to the explanation that linguistic expressions themselves are responsible for creating contexts for accepting a particular inference. If you do not agree, I still hope you might admit that one can hardly expect to find formal (exceptionless) rules in ordinary language when it comes to introducing new obligations from given ones – whenever the result of using a purported rule is at odds with some of utterer's purposes, the applicability of this rule is turned down. These rules are *ad hoc* by their very nature and this is the reason I prefer to avoid labelling them as "logical inferences".

Why, then, do we have a valid formal logic in case of statements and not in case of obligations? The diference between the two schemas (10) and (11) may well be the reason. In case of T-schema (11), although it cannot guarantee its consequences to have practical value ((13), no doubt, is useless), it guarantees that any true statement is taken into a true statement, whereas the F-schema (10) guarantees that a fulfilled obligation is taken into a fulfilled obligation. The difference is, as already stated above, that whereas the truth of a possible statement may be taken as an entitlement to make this statement, fulfilling a possible obligation is in no way related to imposing this obligation. Therefore it does not make sense to talk about the existence of actual obligations resulting from formal entailment.

The widespread assumption that, nevertheless, some entailment rules for obligations *must* exist may be supported by an dubious picture of what is going on in case of assertoric utterances. Namely, one may think that the possibility to occasionally replace, in ordinary speech, one statement with another, in a situation where the first statement cannot be true without the second statement being true, depends on the existence of a formal logic where similar replacements are allowed without vague restrictions. I suggest that things may well be the other way round – that substitution practices of ordinary language should be regarded as a primary linguistic device, whereas formal logic is just a kind of extrapolation of allowable substitutions where occasional fading of the original purpose – to facilitate communication – is compensated by the clarity and universality of the application rules. When it comes to obligations, however, no meaningful extrapolation is available.

4. Conditional obligations.

Although F-schema (10) is a useful device to clear up certain problems – or that's what I hope –, it need not cover all kinds of obligations.

Let us start with Dummett's example[9], mother telling her son,

(15) If you go out, put on your coat.

The son, not finding his coat, fulfills her order by not going out. *This* example nicely conforms to (10),

(16) Obligation($p \rightarrow q$) is fulfilled $\equiv p \rightarrow q \equiv \neg p \vee q$.

[9] Dummett 1959: 9.

Consider, however, a slight modification of the previous example, mother telling the son,

(17) If it rains and you go out, put on your coat.

Now, if there happens to be no rain, the son can in no way violate the order. The question is, can he fulfill it? No doubt, someone might insist that, whenever an order is not violated, it is fulfilled – that it is exactly what is meant by the expression 'to fulfill the order'[10]. But having accepted the latter claim, one cannot argue any more that the order is fulfilled due to a deliberate choice of the subject, and my claim is that subject's ability to choose between 'fulfillment' and 'violation' is crucial for 'fulfilling the obligation' to make sense.

Thus I suggest that in cases like (17) we should allow the possibility that an obligation will be neither fulfilled nor violated. (I have been using the word 'violated' instead of 'not fulfilled', to express the idea that the current claim does not depend on rejecting the law of excluded middle for statements, or classical logic in general.) E.g., in case of (17) we could say,

(18) The obligation (17) is fulfilled iff (it rains) and (the son does not go out or puts on his coat),
the obligation (17) is violated iff (it rains) and (the son goes out without putting his coat on),

(Here brackets just mark the application order of logical connectives.) It is already a consequence of (18) that the obligation (17) is neither fulfilled nor violated iff it does not rain.

I happily admit that (18) as a construal of (17) violates "syntacticism", since the roles of 'raining' and 'going out' are distinctively different in (18), although they appear in syntactically interchangeable positions in (17). I believe that making this difference, depending on not just syntax, but on what is and what is not considered to be under the control of the subject, is well motivated.

In case of (17) the indifference-clause depends on a part of the antecedent of an explicitly conditional obligation. However, this conditionality need not always be explicit. Take, e.g., another Dummett's example, about some Jones, now dead, who never in his lifetime was in a dangerous situation and who therefore cannot be said to have been either brave or not-brave[11]. Now, some people disagree with Dummett on grounds that we might be able, some day, to evaluate the appropriate counterfactual – 'if Jones had been in a dangerous situation he would

[10] In fact, Dummett also insists that this is how things are (Dummett 1959: 9). His argument is that, unlike in the case of a condtional bet where money can move in one direction, in the other direction, or not move at all, in case of conditional obligation we have only two possible outcomes, punishment or the lack of punishment.

Well, first it is not that difficult to imagine a prescription for some critical situation, say, in a military environment, such that any action in this situation, if the situation occurred, would lead to either an award or a punishment, thus there being three possible outcomes, just like in case of a conditional bet. Second, if the outcome were the only criterion to identify obligations then the lack of punishment would not allow to discriminate between obedience to an order and the lack of orders. Thus, clearly something else but the outcome should be taken into count to identify obligations, and I can see no reasons why this 'something else' could not discriminate between 'the obligation was fulfilled' and 'obligation was neither fulfilled nor violated'.

[11] Dummett 1959: 15.

have …' – and in this way still find out whether Jones was brave or not. Correct or not, this objection cannot be raised in case of corresponding obligation (considering the situation where Jones was obliged to act bravely), as clearly only actual events count in case of obligations. So,

(19) Jones's obligation to act bravely
is fulfilled if he meets some dangerous situation and he acts bravely in all of them,
is violated if in some dangerous situation he meets he does not act bravely.

Again, as a consequence, the obligation is neither fulfilled nor violated iff Jones, never in his lifetime, meets a dangerous situation.[12]

To summarize, I suggest that there may be obligations that depend on two propositions p and q, to be denoted as O(p/q), where the part p is determined (a) by the fact that the occurrence of p is not under the control of the subject, or (b) neither achieving nor avoiding p is a part of the obligation (as I implicitly assumed in case of (19)),

(20) O(p/q) is the obligation that
is fulfilled iff $p\&q$,
is violated iff $p\&\neg q$.

As a consequence, the obligation O(p/q) is neither fulfilled nor violated iff ¬p. The latter may be called the indifference clause of the obligation, as I have already done.

5. Propositional attitudes (traditionally so called)

When it comes to making inferences in case of propositional attitudes – desire and belief –, we constantly meet situations similar to the ones discussed above with respect to obligations. E.g., it is hard to accept that from my wish to drink coffee new wishes, like a wish to drink tea or coffee, can be deduced. Or, that from my belief that John is guilty anybody could derive my belief that John or Mary is guilty. At the same time, it is plausible to claim that if I do happen to have a desire to drink tea or coffee, it is the very meaning of this desire that either beverage will satisfy me; and if I do believe that John or Mary is guilty then John's guilt, as well as Mary's guilt, will confirm my belief.

To combine the talk about the three kinds of attitudes as I will call them (since, as mentioned in the very beginning, I have not been able to figure out any better word to cover the three – obligations, desires and beliefs), I suggest the following generalization of the F-schema (10)

[12] Someone might insist that Jones, a clerk in a rational society, *has* fulfilled the relevant obligation if he has succeeded to avoid dangerous situations. But as easily we may imagine that some other Jones, a knight in the Middle Ages, has not fulfilled *his* relevant obligation if he has not succeeded in finding dangerous situations. I have to agree, as a result, that things are more complicated than I have presented them so far, or will ever present them. What I still maintain is that if there may exist such conditional obligations that non-occurance of the antecedent counts as fulfilment of the obligation, and such that it counts as its violation, it makes sense to assume that there may also exist conditional obligations where non-occurance of the antecedent counts neither as fulfilment, nor as violation of the obligation. And it is my only claim concerning 'conditional obligations'.

(21) Attitude(q) finds positive solution if and only if q .

I have chosen the predicate 'finds positive solution' as the word 'positive' seems to be a suitable candidate for the task, on one hand, to be usable as a primitive, without further explanations, when we are interested only in general inferential properties and, on the other hand, to be supplementable with necessary qualifications, when we are interested in discriminating between obligations, desires and beliefs. For the latter purpose we can pick up two aspects of 'positive', (1) as a ground for striving for something and (2) as an estimation to a particular situation. Obligations, then, focus on the first aspect, leaving the second open, whereas desires focus on the second. Beliefs combine the two – I will be pleased if my beliefs are confirmed, so I strive to form such beliefs. In case of obligations and desires, either striving or estimation will take place in future, whereas in case of belief, striving has taken place in the past while the estimation is expected to take place in the future. (The above explanation is just a justification for using the word 'positive', not an attempt to formulate a theory of different attitudes.)

Conditional attitude A(p/q), generalization of (20),

(22) A (p/q) is the attitude that
finds positive solution iff $p\&q$,
finds negative solution iff $p\&\neg q$

(and which, consequently, finds neither positive nor negative solution, if not-p), also seems to be a useful device, although in case of beliefs and desires it is somewhat problematic to formulate criteria for deciding whether my particular 'if p then q' attitude has an indifference clause. I do think that I have numerous 'if p then q' beliefs where I definitely do not care about p-cases, considering my belief to be confirmed only if 'p and q', but I have not found any knock-down argument to convince someone who might suggest that it is better not to include such notion among basic concepts. (My raw guess is that the strict borderline between (21) and (22) does not exist, and this is the reason why it is difficult to propose exact criteria to distinguish the two.)

***

I am not aware of any serious objection to construe desires according to (21) – or (22) –, but I do see some problems specific to beliefs that require certain qualifications. I will turn to these problems in Secs. 7 and 8. Meanwhile, in the next section, I would like to show how the current approach helps to clear up things in case of a well-known example about making inferences in case of beliefs.

6. McGee's counterexample to *modus ponens*.

In McGee's example (McGee 1985), just before 1980 US presidential election, one has good reasons to believe, relying on opinion polls, that the winner will be Reagan, and therefore that

(23) A Republican will win the election.

One also believes,

(24) If Republican wins the election, then if it's not Reagan who wins it will be Anderson,

since just one other Republican, Anderson, participates in the race. And nevertheless there is no reason to believe that

(25) If it's not Reagan who wins, it will be Anderson,

since Carter, a Democrat, has much better prospects than Anderson.

The example was presented as a counterexample to *modus ponens*, but I will start with an alternative way one could infer (25) from (23), namely via

(26) Reagan or Anderson will win.

As far as the point of the example does not depend on there being other Republicans on stage, I assume there are none, and then (23) will be materially equivalent to (26); (26), in turn, is logically equivalent to (25), so (24) and (26) are equivalent to each other. Therefore, arriving at problematic (25) does not depend on using complicated devices like embedded conditionals in (24), we get the same outcome relying on two simple equivalence relations.

Certainly, logical equivalence between (26) and (25) depends on interpreting (25) as a material implication, and it is almost a received view that the latter fails to grasp something essential expressed by the 'if … then …'-construction in an ordinary language. So, as soon as we blunder into a counter-intuitive result involving 'if … then …', it will be the first one to be accused of bringing about the confusion and the one whose customary properties will come under attack.

Now, I think it is not reasonable to suppose that we are always dealing with a universally valid logic and just some problematic connectives like, say, a strong conditional, added to it. Rather, it is the character of the whole logic that matters, and right now we are dealing with the logic of beliefs[13] where, I claim, an important motive for entertaining disjunctive beliefs is the possibility to rely on the rule,

(27) Whenever I believe that p or q, I also believe that if not p then q ,

whatever interpretation we might give to the 'if … then'-construction.[14] With the two interpretations considered so far, believing material implication, an instance of $A(\neg p \rightarrow q)$, will be simply equivalent to $A(p \vee q)$, whereas the belief with the indifference clause, an instance of $A(\neg p/q)$, would express something like "it the p-option happens to drop out, what

[13] Somebody might claim that McGee's example can be stated independently of beliefs, in terms of objective expectations. Even if it were so, i.e. if objective expectations were not reducible to, say, well-grounded beliefs, we would be still in position to rely on schema (21), since I cannot see any reason to deny its ability to deal with expectations, whatever their true nature. And possibility to rely on (21) is all that matters in the current argument.

[14] The last claim reminds one of the form 'Belief(p or q) entails Belief(if not p then q)', and as such, seems to contradict to the main claim of this paper that beliefs, like attitudes in general, do not follow from each other.

This question, I admit, would require a further analysis I am not ready to offer at the moment; as a preliminary answer I suggest that the case at hand is somewhat specific in that the second belief is a part of the meaning of the first, where the meaning of a linguistic device depends on the purpose of its use.

remains is to believe that *q*". And if there were some other interpretation of 'if-then"-belief available, I still cannot imagine it could violate (27).

If the above claim is true then the objectionable (25) can be rejected only together with (26), and now the problem is that the latter appeared to be correct at the first sight. Nevertheless I think that rejecting (26) is a reasonable thing to do.

To start with, let us consider a general case of excluding each other three expectations x_1, x_2, x_3 where x_1 is the likeliest outcome, x_2 its closest rival and x_3 a distant third. A natural way to describe my expectations is to say that (1) I believe that x_1, (2) I believe that if not x_1 then x_2 and (3) I believe that if neither x_1 nor x_2, then x_3. Now, the second and third can be interpreted as conditional beliefs $A(\neg x_1 / x_2)$ and $A(\neg x_1 \,\&\, \neg x_2 / x_3)$, respectively (notation defined by (22) has been used here), but they can also be interpreted as containing material implication, $A(\neg x_1 \rightarrow x_2)$ and $A(\neg x_1 \,\&\, \neg x_2 \rightarrow x_3)$. The latter two are, of course, equivalent to $A(x_1 \vee x_2)$ and $A(x_1 \vee x_2 \vee x_3)$. That is, although I have the belief that x_1, I may also be said to have the belief that x_1 or x_2 (the one that is less likely to fail than my primary belief) and then also the belief that x_1 or x_2 or x_3 (where the possibility to fail is almost ruled out). Yet I do *not* have the belief that x_1 or x_3. Why should I? Why should I believe that Reagan or Anderson will win?

Certainly, it is not unusual to claim that one can apply the rule "since a disjunction is true if at least one of its disjuncts is" to deduce 'Reagan wins or Anderson wins' from 'Reagan wins' (e.g., Katz 1999:411), ignoring the fact that we are talking about beliefs (or expectations) here. What may encourage this is, I guess, the frequently repeated claim that there is no difference between believing that *p* and believing that *p* is true. My point is that this claim is erroneous: believing that something is true amounts to believing that the corresponding statement would be true, and we can infer new statements from existing ones, whereas we cannot infer new beliefs. From the statement 'Reagan wins' the statement 'Reagan or Julius Caesar wins' follows (or, more precisely, truthfulness of the latter statement follows from truthfulness of the former), but there is no ground for similar inference in case of beliefs. I may believe that 'Reagan or Julius Caesar will win' is true, without believing that Reagan or Julius Caesar will win. (My accepting the former belief may still depend on some training performed in logic classes.)

Now, do I believe (23), that a republican will win? I do, no doubt, if what is meant is the belief that the person who wins will be a Republican; yet I do *not* if what is meant is the belief that arbitrary person from the set of Republican candidates will win. Just like from my obligation to pass the examination on next Thursday it may seem to follow that I have the obligation to pass it next week, from my belief 'Reagan will win' the belief 'a Republican will win' may seem to follow; and just like in case of obligations, this apparent inference may be explained rather as applying a pragmatic rule that allows me to use a less specific predicate than the one included in my actual belief if it is all that matters in the given context, together with the assumption that I will use the phrase 'a Republican will win' only in appropriate contexts.

Once the plausibility of the belief (23) has been turned down, McGee's example ceases to be a counter-example to *modus ponens*. Still it does not mean that we should accept *modus ponens* as a device that allows to deduce new attitudes on the basis of original ones: existing attitudes can be in conflict with each other and, in particular, I may occasionally entertain simultaneously three attitudes $A(p)$, $A(p \rightarrow q)$ and A(not-*q*) instead of $A(q)$). However, they will then be conflicting attitudes – the ones that cannot find positive solution simultaneously,

and I will be aware of this fact. In case of beliefs it will mean that I know: at least one of them is wrong and this knowledge compels me to find ways to disbelieve one of them. McGee, on the other hand, seems to claim that he has found an example where *modus ponens* is violated by coherent beliefs. My claim is that he has not, as far as my belief that a Republican will win is properly construed as a belief that a particular Republican will win, whereas to form *modus ponens* a different belief, the belief that some unspecified Republican will win, is needed; and there is no ground for the latter belief.

7. Belief and knowledge.

In case of beliefs, a general objection 'But I *do* infer beliefs from each other' may sound more convincing than in case of other attitudes. As a reply, I suggest to regard beliefs as coming in three stages where only the 2nd should be regarded as an attitude in the sense of the current paper and, as such, a subject to the claim that beliefs do not follow from each other.

The first stage of believing is making statements not taken as expressing personal opinion yet[15], and here the customary logic of statements applies. The phrase 'I believe that …' is usually not used in this case; yet, Jones may convey Smiths statement that *p* as, "Jones believes that *p*"; what was a statement for Jones has become, for Smith, Jones's belief. What distinguishes the latter is explicit awareness of possible error, however little the chance for it is taken to be. As such, it may found either "positive solution" or "negative solution", and it therefore conforms with the general schema for an attitude, (21) or (22).

The awareness of a possible error yields sometimes a want to exclude this possibility, and so we get the third stage – knowledge –, that may be characterized as a belief that pretends to be not (just) a belief. At this point usually talk about 'justified belief' begins, but I suggest rather to talk about a belief about justification,

(28) *s* may say sincerely, "I know that *p*"
if and only if
s believes, "I have a hold over justification that grants that *p*".

Now, a likely objection to (28) will be, "It is in no way a definition of '*s* knows that *p*' since it does not even grant the truthfulness of *p*." But what could grant it? A customary discussion about knowledge is held as if from a point of view of some omniscient person who simply tells me how things are. Or, alternatively, his "description" should be understood as a performative act, in full analogy to 'I name this ship the "Queen Elizabeth" ', as giving rise to the situation he descibes. In either case I am not expected to question the story I am told – Jill drives a Pontiac once they say so. It is very different from deciding whether to use 'I know' in everyday life where somebody just saying something cannot grant truthfulness of any proposition. And in fact – nothing can. All one can do is just to add an additional reference

[15] 'The first' and 'the second' are meant to be not just arbitrary labels, but to express a kind of metaphysical priority – I think that an argumentation similar to the one Sellars has presented to show that "seeing that x, over there, is red" is prior to "it is looking to one as though x, over there, were red", since the former expresses the common descriptive content of the two experiences (Sellars1963: 151), can be used to show that in case of 'Asserting that *p*' and 'Believing that *p*', it is the former that expresses the common content of the two, so that the ability to believe that *p* may be said to depend on the ability to assert that *p*. Similarly, the third stage, the ability to know (not *just* believe) is meant to depend on the ability to believe.

point to the previous 'justification that grants'-formulation, arriving at something like,

(29) s_1 may say at t_1 sincerely that s_2 knows at t_2 that p
if and only if
s_1 believes at t_1 that s_2 has at t_2 a hold over justification that grants that p.

Or, (28) might be extrapolated differently, like

(30) s_1 may say at t_1 sincerely that s_2 knows at t_2 that p
if and only if
s_1 believes at t_1 that s_2 believes at t_2 that s_2 has at t_2 a hold over justification that grants that p.

(In both cases s_1 and s_2 may denote the same person or two different persons, just like t_1 and t_2 may denote the same moment of time or two different moments; the verbs 'says', 'knows', 'believes' and 'has a hold over justification' have been used in a non-grammatical, tenseless and personless way, expressing predicates that depend on a particular moment of time – past, present or future; 'grants' is assumed to be independent of time.)

The difference between the two extrapolations of (28) is, first of all, that while (30) allows me to say 'Yesterday I knew that p, but now I know that not-p', (29) will force me to express myself differently, e.g. like, 'Yesterday I believed I knew that p, but ...'. Besides, (29) conforms to philosopher's intuitions about Gettier's cases, so it is (29) that, in every respect, suits for a philosopher, although I think that many other people use the word 'know' as if it was governed by (30).[16]

Now, although knowledge is defined via belief that was claimed to lack usual inference patterns, on the basis of either (29) or (30) we get back allowance to perform any customary inferring. In particular,

(31) $(p \rightarrow q) \rightarrow (\mathrm{K}(p) \rightarrow \mathrm{K}(q))$,

will be a valid implication if we assume that if p implies q then anything that grants that p will also grant that q; and I see no reason to doubt in this assumption. E.g., if something grants that p, it also grants that p or q, and thus knowledge that p entails knowledge that p or q, in the current approach, just like in a customary approaches.

So, in case of not-yet-beliefs (statements) and not-just-beliefs (knowledge) customary inference patterns remain valid; the claim that we cannot deduce new beliefs from existing ones concerns only a species of attitudes between them.

---

[16] The point of Gettier's examples is to consider justified true beliefs where the actual justification fails, although the belief remains true and may be therefore still described as a 'justified true belief'; according to (29), I cannot call something knowledge if the believed justification has failed – I cannot believe that 'the justification grants', once I am aware that it has actually failed to do so.

The experiments described, e.g., in Weinberg et al 2001, have shown that general public does not see Gettier cases eye to eye – not everybody thinks that they do not qualify as knowledge. At the same time, I am not aware of any phfilosopher challenging Gettier, at least not before the results of the above mentioned experiments have become available. It is this fact that has motivated me to talk about "philosopher's intuition" concerning Gettier's examples.

8. An attempt to cope with opacity of belief contexts.

My main reason for rejecting deducibility of attitudes has been motivated by the schema (21) (or (22)) that allows to regard Attitude(q) as a function that maps q-worlds to the value POSITIVE and not-q-worlds to the value not-POSITIVE, or Attitude(p/q) as a function whose domain is the set of p-worlds that maps (p&q)-worlds to POSITIVE and (p&$\neg q$)-worlds to NEGTIVE). As far as functions do not follow from each other, attitudes do not follow either. However, if we have an equivalence relation $p \equiv q$, the two pairs (p-worlds, not-p-worlds) and (q-worlds, not-q-worlds) do not differ from each other, and so, in particular, Belief(p) and Belief(q) obtain one and the same interpretation. Yet everybody knows that belief context is opaque, i.e. that one may believe a proposition if it is presented in one way, and not believe it if it is presented in some other way. To deal with this possibility, we have to reinterpret the P-schema (21) in a different manner.

When I believe something, my ambition seems to be that the (assumed) content of my belief should not depend on a particular presentation or formulation of what I believe. Yet I am well aware that I may easily fail like, e.g., Kripke's Pierre (Kripke 1979) who, after visiting London, becomes to believe, "London n'est pas joli", while also keeping his previous belief, formed after reading a book in French, "Londres est joli". The problem of specifying proper object of beliefs cannot be solved just by sticking to linguistic forms, as illustrated by Kripke's next example about Peter who believes that Paderewski, a pianist whose concert he has visited, has musical talent, whereas Paderewski, the Polish prime minister in 1919, has not, since, as he believes, politicians never have. (Actually it is the same Paderewski, of course.) Including sense in the meaning of proper names does not solve the general problem either, since it is not difficult to reformulate Pierre's example without using any proper names at all, as shown by the following variant of the previous story.

Imagine Pierre transported, in a dark night, to a foreign city. There he meets Peter who tells him how unlucky they are that it is so dark, as the place is very beautiful in daylight. Peter sounds convincing, so Pierre becomes to believe, "the place where I am now is beautiful". Pierre then goes on a train and falls asleep, while the train remains at rest for the whole night. Waking up the next morning, Pierre mistakenly thinks he has been moved into a new location; he looks out and sees an ugly place, so he becomes to believe, "the place where I am now is not beautiful", while there is no reason for him to give up his previous belief about the same place. As a result, he has simultaneously two contradicting to each other beliefs, just like the original Pierre had, whereas proper names have played no role in the story.

What then makes Pierre's two beliefs different? Perhaps the different reason for believing in each case, being told *vs.* seeing with one's own eyes? But then, does my belief that there is a river in London depend on whether I have seen the river or just read about it? And, on the other hand, might it not happen that I form conflicting beliefs about the same thing, not realizing that it is the same thing, relying both times on what I see?

One could go on with similar examples for quite a while, yet I dare to suggest already on the given basis that whatever criteria one introduces in order to get a more fine-grained representation of beliefs, the resulting picture will still lack necessary details in some situation and will introduce unwanted distinctions elsewhere; I even think that any qualification for belief-contents beyond Russellian propositions will be accepted only when there is a real need

for it – e.g. in order to explain how a rational person can entertain two contradicting to each other beliefs without feeling any tension between the two.

Such *ad hoc* character of acceptable qualifications to belief contents may be explained by a hypothesis that beliefs are the more useful the less they depend on such contingencies like a particular representation of their content, or a particular reason to entertain them; although such dependencies cannot be avoided, it is reasonable to make their role as insignificant as possible, and to admit them only when there is a real need for them, not in advance according to some pre-fixed criteria.

If this is true, and I still want to explain how inferences are made in case of beliefs with the help of something similar to the P-schema (21), but counting also for possible opacity, I have to specify the metalinguistic *q* in (21) as something whose identity criteria are as loose and *ad hoc* as those of beliefs, and I propose that 'evidence that *q*' may do the job. That is, I hope that nobody is ready to give criteria that would allow to determine with a certainty when two instances of 'evidence that *q*' should be considered the same evidence, and when not, since primarily it is 'what is this evidence for' that matters, whereas 'what does this evidence rely on' will be added to the extent it is needed.

Now, if we just rewrote (21) using 'evidence that *q*' instead of *q* on the right side, then, about possible failure, all we would be able to say is, 'If there is no evidence that *q* then Attitude(*q*) will not find positive solution'. However, whereas it makes sense to assume that any world that is not a *q*-world is a not-*q*-world (many people take the law of excluded middle for granted, anyway), it would be quite implausible to assume, even for a while, that the lack of evidence that *q* might generally imply the evidence that not-*q*. Thus, 'evidence that not-*q*' will not appear automatically in the schema, so if I want it to be there – and I do –, it must be added separately. These considerations lead to the following formulation,

(32) Attitude(*q*) will find positive solution iff there is evidence that *q*, and
Attitude(*q*) will find negative solution iff there is evidence that not-*q*.

If there is neither evidence that *q* nor evidence that not-*q* (no *q*-related evidence), the Attitude(*q*) will find neither positive nor negative solution (will find no solution).

Although the purpose of (32) was to count for episodic differences of materially equivalent belief contents, it also introduces at least one kind of persistent distinction, depending on the type of quantification for the evidence.

Consider a belief about a particular species of a bird xxaborra, 'All xxaborras are green', shared by two English speaking persons, Jones and Smith, whose beliefs differ in the following respect: Jones's belief is based on his accidental reading of an ornithology book where he met the statement 'All xxaborras are green' that made him believe that only green birds are called xxaborras. On the other hand Smith, an experienced ornithologist, knows that identifying a bird as a xxaborra does not depend on its color; still, all xxaborras he has seen were green, and he has never heard about any other specialist meeting a non-green xxaborra either. What I suggest is to distinguish between the two versions of the belief 'All xxaborras are green', calling one of them 'universal belief' ("Any xxaborra I happen to see will be green") and the other 'existential belief' ("There exists an agreement among ornithologists according to which only green birds are called xxaborras").

To construe 'a whale is a mammal' as 'every *x* is such that if it is a whale then it is a mammal' was Frege's ingenious insight, since only such construal provides clear truth conditions for this kind of statements. Still it has always disturbed me that, as a result, a claim about an agreement on classification and an empirical claim like 'all ravens are black' both acquire the same logical form. According to the current suggestion they will have different form when it comes to corresponding beliefs – the first will become an existential belief ('there exists an agreement that...') whereas the second remains universal ('every raven is ...'). On these lines, any mathematical sentence of the form 'every *F* is *G*' will be coupled to an existential belief, 'there exists a proof showing that if something is *F* then it is *G*'. While 'every *F* is *G*' may be true even if there are no *F*s, my believing that every *F* is *G* always relies on there being something (e.g., a proof or an agreement), therefore conforming to (32) in that any solution of a belief requires the availability of some evidence.

The two sentences, 'If there existed the greatest prime *p* then *p*! + 1 would be a prime' and 'If there existed the greatest prime *p* then pigs would fly' may be both true, as suggested by David Lewis (Lewis 1973), but it is only the first sentence of the two that I believe.

9. The basic biconditional: about possible asymmetry between the two conditionals.

In Section 3 I suggested that a strong feeling as if we had a valid instance of logical reasoning in the little-brother case (1), is misleading and that relying on a couple of pragmatic rules could explain this feeling. I was motivated by taking seriously the proposed F-schema which does not give reason to deduce new obligations from existing ones – telling truth just to one's little brother does not entail there being an additional obligation one fulfills, it just avoids violation of the original obligation.

But what if the two conditionals forming biconditional in (10) are not of equal importance? Perhaps one of them may be ignored, and Obligation(*p*) just means, roughly, something like 'if not-*p* then punishment'? Since, one may argue, the other conditional would be 'if *p* then award', and punishing for bad behavior is much more common in human practice than awarding for just good behavior, so the second conditional may be practically ignored. And if so, the pattern (3) becomes a valid form of logical reasoning.[17]

Now, the need to rely on a conceivably contingent practice to ground a purportedly logical relation would hardly please a logician. Besides, is it sure that the existing practice rather to punish than to award helps to explain the inference in case of (1)? Has anybody heard of somebody being punished for not telling truth to his little brother? More over, I think that despite the existing practice we should be able to imagine a custom never to punish for lying, but to award those who have told only truth during their lives (say, their names would be engraved in marble); it seems that even such a scenario does not weaken the intuition that one is entitled to make inferences like (1), although now telling truth to just one's little brother clearly does not lead to any consequence.

Yet there is another kind of salient asymmetry between the fulfillment and non-fulfillment parts of (1), namely the type of quantification needed in either case: in order to fulfill the

[17] Assuming, 'Obligation(*p*) $=_{def} \neg p \rightarrow$ Punishment', the relation '$p \rightarrow q$' gives rise to the following succession: '$\neg q \rightarrow \neg p$' (contraposition of the original relation), '$\neg q \rightarrow$ Punishment' (Obligation(*p*) and transitivity of implication), 'Obligation(*q*)' (the purported definition).

obligation one has to tell truth in every situation where the 'telling truth' predicate is applicable, whereas lying just once qualifies as its violation. And this concerns, of course, any obligation (and attitude in general) that contains specifications like 'to anyone' or 'always' (the specification need not be explicit), i.e. a universal quantification over relevant situations. One might suggest, then, that just like laws of nature can be only falsified and never verified, only the not-fulfilled part is (practically) available for us, whereas the possibility to evaluate every single relevant act in the life of a particular person is not; and therefore it is primarily the 'not-fulfilled'-part that matters.

Well, I agree that in case of universal obligations the not-fulfilled part is more practical,[18] still it will not finally solve the problem with the intuitive validity of (1)[19]. Since when I want to make up my mind whether somebody follows (1) or not, it will not suffice that I have never heard him lying, I also need to hear him telling truth, preferably in situations where people are often tempted to lie; the actual instances of fulfillment matter no less than the lack of instances of violation. In general, the following principle seems to hold,

(33) If I must behave *Y*-ly in any situation of type *X*, and *x* is a situation of type *X*, then I must behave *Y*-ly in *x*.

I accept (33) as a valid instance of a rule governing the usage of the word 'any',

(34) If I am entitled to say *Y* about any *X*, and *x* is *X*, then I am entitled to say *Y* about *x*.

The rule (34) is a vital part of the meaning of the word 'any' (the main purpose of using it), still I would not call it 'a logical rule', since I think it has, if not clear exceptions, at least controversial instances, like those that introduce existential import; e.g.:

(35) If there is an obligation concerning any situation of type *X*, and *x* is a situation of type *X*, then there is an obligation concerning *x*.

What we are dealing with in case of (35) is something like a converse problem of universals: whereas the problem of universals concerns questions like, "Does the fact that there is particular redness in every red thing entail the existence of a universal redness apart from particular things?", the converse problem concerns questions like, "Does the existence of a general obligation entail the existence of a particular obligation in a situation where the general obligation applies?". The idea here is that in both cases we have, first, an ordinary way to exist and then, perhaps, some exotic way. In case of redness, ordinary existence occurs in red things and in case of obligations ordinary existence depends on the obligation having

[18] The overwhelming tendency to formulate universal obligations in negative – "Thou shalt not lie" instead of "Thou shalt always tell truth" – may be related to this asymmetry of practical availability of fulfilled and not-fulfilled parts of the basic biconditional: if the obligation is formulated in negative the relevant predicate like 'lie' enters the not-fulfilled in its original form (the norm is not fulfilled if I lie on an occasion) whereas in case of (1) the predicate 'telling truth' has to be negated (the norm is not fulfilled if I do not tell truth on an occasion). Probably it makes sense to keep the practically better available part more straightforward.

[19] I admit that there is no reason to rule out the possibility of there being, say, prohibitions that truly lack the fulfilled-part and are characterized by the not-fulfilled part alone, although it is not that easy to say when exactly one has to give up the fulfilled-part in (10). I still think, however, that in case of a general analysis of obligations it is reasonable to keep both conditionals; it then depends on a particular situation which one of the two conditionals will acquire greater significance.

been imposed somehow – a parliament has approved a law, a tradition has given rise to a norm, a person has given a promise. Assuming the existence of obligations that have not come into being in the ordinary way – by being imposed – is then analogous to supposing the existence of redness apart from red things, as both assumptions lead to a different from ordinary kind of existence.

If I am entitled to think that propositional attitudes are related to certain states of mind, the exoticism of the existence of deduced propositional attitudes is even more salient. I believe, say, that all humans are mortal and I assume there is some state of mind corresponding to this belief. If I also had a corresponding belief concerning every particular human I can think of, there would be thousands of additional beliefs and thousands of additional states of mind. Such a picture would lead to unnecessary complications, I think.

***

Together with (33) I have accepted a variant of (1),

(36) I must tell truth to anybody. Therefore, I must tell truth to my little brother.

Now, if someone thinks that (1) is synonymous to (36) and accepts (1) on this basis, my disagreement with him is rather insignificant: he most likely takes 'there is …' in (1) not too seriously and on such conditions I am ready to accept (1) without further reservations. On the other hand, if someone acknowledges the difference between (1) and (36), but still accepts (1) as a valid instance of reasoning, I have my reservations described a few paragraphs earlier – that then the existence expressed by 'there is …' should be understood as some exotic kind of existence. In either cases, however, my main point is that neither (1) nor (36) is a distinctively deontic kind of reasoning, but just an instance of a general rule governing the usage of words that express universal quantification in ordinary language. And that it is not obvious that one is entitled to call this rule logical, since it has instances that are, if not invalid, at least controversial.

10. The role of logic in case of attitudes.

Juliano Maranhão (Maranhão 2009) has suggested that we have two alternatives when talking about the possible role of logic in case of norms:

> Suppose that in an imaginary country every male has the duty either to join the army or to do charity work when he completes 18 years old. Suppose now that a new law suppresses the army in that country. If the congress enacts a new law stating that "males ought to do charity work at the age of 18" does it make any difference to the existing normative order? If it does then logic plays no role in the ontology of norms; if it does not, then logic is a sufficient condition of existence of a norm.

My suggestion is that either answer to the question – 'yes' and 'no' – is plausible, depending on the reading of 'the existing normative order'. If it is construed as a set of existing norms then the answer is, according to the approach of the current paper, 'yes, it makes a difference'; if it is construed as a range of available options for a citizen obedient to the laws then the answer is, 'it makes no difference'. Accordingly, I do not agree to Maranhão's claim that the answer 'no' should lead to the conclusion that logic is a sufficient condition for the *existence* of a norm – the role of logic here is not to deduce new norms, but just to find out alternatives an obedient to laws citizen has. In this, I agree to von Wright's claim that deontic logic can tell us nothing about the existence of norms.

This observation does not concern just deontic concepts; the conclusion I would like to finish with is that the role of making inferences in case of any kind of attitudes – obligations, desires, beliefs – is not to deduce the existence of new attitudes, but just to make decisions about different situations, both actual and possible, to realize whether, in a given situation, a particular attitude finds positive solution, negative solution, or no solution.

References

Chisholm, Roderick M. 1963. Contrary-to-Duty Imperatives and Deontic Logic. — *Analysis* Vol 24, pp 33-36

Davidson, Donald 1977. The Method of Truth in Metaphysics. — *Inquiries into Truth and Interpretation.* Oxford 2001, pp 199-214.

Dummett, Michael 1959. Truth. — *Truth and Other Enigmas.* Duckworth 1992, pp 1-24.

Katz, Bernard D. 1999, On a Supposed Counterexample to Modus Ponens. — *The Journal of Philosophy*, Vol. 96, No. 8, pp 404-415.

Kripke, Saul 1979. A puzzle about belief. — *Meaning and use*, A. Margalit (ed.), Reidel, Dordrecht, pp 239-283

Lewis, David 1973. *Counterfactuals*. Harvard University Press.

Maranhão, Juliano S.A. 2009, Von Wright's Therapy to Jørgensen's syndrome. — *Law and Philosophy*, Vol 28, pp 163–201.

McGee, Vann 1985. A Counterexample to Modus Ponens. — *The Journal of Philosophy,* Vol. 82, No 9, pp 462-471.

Ross, Alf 1941. Imperatives and Logic, *Theoria*, vol 7, pp 53-71.

Sellars, Wilfrid 1963. Empiricism and the Philosophy of Mind. — *Science, Perception and Reality*. Routledge & Kegan Paul 1963, pp 127-196.

Weinberg, Jonathan M., Nichols, Shaun and Stich, Stephen 2001. Normativity and epistemic intuitions. — *Philosophical Topics* vol. 29: 429–460.

Wright, Georg Henrik von 1991. Is there a Logic of Norms? — *Six Essays on Philosophical Logic*. *Acta Philosophica Fennica* Vol 60 (1996), pp 35-54.

Åqvist, Lennart 1967. Good Samaritans, Contrary-to-Duty Imperatives, and Epistemic Obligations. — *Noûs*, vol 1, pp 361–379.

Anto Unt, Tartu University.
May 17, 2010.

Resümee

Käesolevasse magistritöösse on koondatud kolm artiklit, millest kaks on avaldatud, kolmas mitte. Artikleid ühendab idee käsitleda kohustusi, lähtudes skeemist "Kohustus O(*p*) on täidetud parajasti siis, kui *p*". Võrreldes seda skeemi Tarski T-skeemiga, näidatakse kolmandas artiklis, miks väidete puhul saab rääkida nende üksteisest järeldumisest, kohustuste puhul aga mitte. Kohustuste kohta käivat käsitlust üldistatakse ka propositsioonilistele hoiakutele – uskumustele ja soovidele; üldisel juhul, mis hõlmab nii hoiakuid kui ka kohustusi, võetakse aluseks skeem "Hoiak A(*p*) leiab positiivse lahenduse parajasti siis, kui *p*." Näidatakse, kuidas selline lähenemine võimaldab heita uut valgust McGee *modus ponens*'it kahtluse alla panevale näitele ja kuidas määratleda teadmist nii, et see oleks kooskõlas Gettieri näidetega.