

Ville Pehkonen

**AUTONOMISET AJONEUVOT JA ETIIKKA:
PÄÄTÖKSENTEKO ONNETTOMUUSTILANTEISSA**

Informaatioteknologian ja viestinnän tiedekunta

Kandidaattitutkielma

Toukokuu 2020

TIIVISTELMÄ

Ville Pehkonen: Autonomiset ajoneuvot ja etiikka: päätöksenteko onnettomuustilanteissa
Kandidaattitutkielma
Tampereen yliopisto
Tietojenkäsittelytieteiden tutkinto-ohjelma
Toukokuu 2020

Tutkielman tavoite on perehtyä autonomisen ajoneuvon päätöksentekoprosessiin onnettomuustilanteissa ja selvittää minkälaisia eettisiä seurauksia päätöksenteon perustana toimivien algoritmien suunnittelulla on. Tutkielma toteutettiin kirjallisuuskatsauksena. Aineisto koostuu pääasiassa vertaisarvioituista tutkimuksista, konferenssipapereista ja kirjoista.

Autonomisella ajoneuvolla tarkoitetaan ajoneuvoa, joka pystyy liikkumaan paikasta toiseen ilman ihmisen avustusta. Vaikka autonominen ajoneuvo pystyisi havaitsemaan ympäristön täydellisesti ja toimimaan turvallisemmin kuin ihminen aiheutuu ennen pitkää tilanne missä autonominen ajoneuvo joutuu onnettomuuteen. Onnettomuustilanteissa toiminta tulee käsitellä ohjelmallisesti. Autonomisen ajoneuvon tulee minimoida vahingot onnettomuudessa sekä tienkäyttäjien fyysisten vammojen, ajoneuvon ja ympäristön aineellisten vahinkojen, että eläinten vammojen suhteen. Optimaalisen liikeradan valitsemiseen ei ole yksinkertaisesti toteutettavaa vaihtoehtoa.

Päätöksentekoon onnettomuustilanteessa tutustutaan esittämällä autonomisten ajoneuvojen etiikkaan liittyvässä kirjallisuudessa yleinen raitiovaunuongelma. Raitiovaunuongelmalla tarkoitetaan tilannetta, missä autonomisella ajoneuvolla on kaksi tai useampi mahdollinen liikerata tilanteessa, missä kaikki vaihtoehdot johtavat onnettomuuteen. Tutkielman aineistossa nousi esiin kolme erilaista teemaa, miten päätös liikeradasta tehdään. Kolme teemaa ovat seurausetiikan alle lukeutuva hyötyetiikka ja sen rinnastaminen kustannusfunktioihin, velvollisuusetiikka ja sen vertaaminen sääntöjen noudattamiseen, ja viimeiseksi nykypäivänä ja lähitulevaisuudessa autonomisissa autoissa käytetyt menetelmät. Kaksi ensimmäistä teemaa vaativat toimiakseen luotettavuutta, mikä ei ole nykypäivänä tai lähitulevaisuudessa ole saavutettavissa. Tämän hetken suurimmat ongelmat ovat kohteiden havainnointi, luokittelu ja lokalisointi. Sen sijaan että kehitettäisiin vain sensoreita ja tiedonkäsittelyä, pidetään lähitulevaisuudessa tärkeänä kommunikaation lisääntymistä ajoneuvosta toisiin ajoneuvoihin, muihin tienkäyttäjiin ja ympäristöön.

Avainsanat: Autonominen ajoneuvo, etiikka, liikenneturvallisuus, päätöksenteko,

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla

Sisällysluettelo

1. Johdanto	1
2. Autonominen ajoneuvo	2
3. Raitiovaunuongelma	3
4. Päätöksenteko onnettomuustilanteissa	4
4.1 Seurauseettinen lähestyminen	4
4.2 Velvollisuuseettinen lähestyminen	6
4.3 Nykytilanne ja lähitulevaisuus	7
5. Keskustelu	8
Lähdeluettelo	9
Kuvien lähteet	11

1. Johdanto

Itseohjautuvalla eli autonomisella ajoneuvolla tarkoitetaan ajoneuvoa, mikä pystyy havainnoimaan ympäristöään ja liikkumaan paikasta toiseen ilman ihmisen avustusta. Autonomisten ajoneuvojen yksi suurimpia myyntipuheita ovat turvallisuuden ja liikkuvuuden parantaminen. Toisin kuin ihmisen, automatisoitujen ajoneuvon keskittymisky ei herpaannu kännykän soimisesta tai radiokanavan vaihtamisesta. Tietokone ei koe väsymystä tai tylsyyttä. Vaikka automatisoidussa ajoneuvossa olisi täydellisesti ympäristöä havainnoivat sensorit ja muiden liikenteessä olevien liikkeitä ennakoivat ohjelmistot, eivät ajoneuvot kykene maantienopeuksissa pysähtymään riittävän nopeasti estääkseen jokaista liikenneonnettomuutta (Goodall, 2014a). Ajoneuvojen suuri massa ja nopeus tieliikenteessä vaativat jatkuvaa riskianalyysiä, eikä onnettomuuksilta voida välttyä. Erilaisten automatisoitujen ajoneuvojen määrä tieliikenteessä on kasvanut viime vuosina valtavasti, minkä myötä myös onnettomuuksien lukumäärä on kasvanut. Esimerkiksi Kaliforniassa tapahtui koko vuonna 2015 yhdeksän kolaria joissa oli toisena osapuolena autonominen ajoneuvo. Pelkästään tammikuussa 2020 vastaava lukumäärä oli 13 (DMV, 2020).

Osana muuta liikennettä automatisoidut ajoneuvot tulevat tekemisiin muun muassa ihmisten ajamien ajoneuvojen, jalankulkijoiden ja toisten automatisoitujen ajoneuvojen kanssa. Pyrkinessään saavuttamaan määränpänsä automatisoidut ajoneuvot priorisoivat erilaisia tavoitteita jatkuvasti. Näitä voivat olla turvallisuus, polttoainetehokkuus tai määränpänsä saavuttaminen mahdollisimman nopeasti (van Loon ja kumppanit, 2015). Tavoitteiden painoarvoja muuttamalla voivat autonomiset ajoneuvot toimia eri tavoilla samankaltaisissa tilanteissa. Toiminnan kehyksinä ovat teknologiset rajoitteet, muu liikenne ja tieliikennelait. Kehyksen sisällä toimiminen vaatii sekä sääntöjen noudattamista että kykyä sopeutua niin tavanomaisiin kuin poikkeuksellisiin tilanteisiin.

Tässä tutkielmassa perehdytään mihin automatisoitujen ajoneuvojen päätöksenteko perustuu onnettomuustilanteissa ja mitä eettisiä seurauksia päätöksenteon perustana toimivien algoritmien suunnittelulla on. Tämä tehdään kirjallisuuskatsauksena alan tieteelliseen kirjallisuuteen ja julkaisuihin. Julkaisuissa nousee esille kolme teemaa, joita käsitellään tarkemmin. Näistä ensimmäinen on seurausetiikan (*Consequentialism*) ja erityisesti tämän alle lukeutuva hyötyetiikan (*Utilitarianism*) vertaaminen kustannusfunktioihin (*cost functions*) ja jonkinlaisen hyödyn maksimoimiseen. Toinen on velvollisuusetiikan (*Deontology*) vertaaminen sääntöjen noudattamiseen. Nämä kaksi teemaa edellyttävät ajoneuvon sensorien havaitsevan ja tunnistavan kohteita tarkemmin kuin mihin nykyteknologia kykenee. Kolmannessa teemassa selvitetään mitä nykypäivän ja lähitulevaisuuden suurimmat teknologiset rajoitteet ovat ja minkälaisia eettisiä ongelmia käytetyt menetelmät aiheuttavat. Lopuksi käydään keskustelua miten ja miksi suhtautuminen näihin kolmeen teemaan on muuttunut automatisoidun ajamisen lyhyen historian aikana.

2. Autonominen ajoneuvo

Autonomisella ajoneuvolla tarkoitetaan tässä tutkielmassa moottoroitua ajoneuvoa, mikä pystyy kulkemaan paikasta toiseen ilman ihmisen avustusta. Society of Automotive Engineers (SAE) määrittelee automatisoidulle ajolle kuusi tasoa, nolasta viiteen. Ajoneuvossa voi olla ominaisuuksia jotka sijoittuvat eri tasoille. Tarkastelemalla vakionopeudensäätimen eri toteutuksia voidaan käydä läpi eri tasojen eroja. Tasolle nolla sijoittuu vakionopeudensäädin mikä ainoastaan pitää yllä käyttäjän valitseman nopeuden. Tällainen vakionopeudensäädin ei ota huomioon ympäristöään millään tavalla. Jos vakionopeudensäädin mukauttaa nopeutta edellä ajavan ajoneuvon liikkumisen perusteella, mutta kuljettaja hallitsee auton ohjausta lukeutuu vakionopeudensäädin tasolle yksi. Jos vakionopeudensäädin hallitsee auton ohjausta nopeuden lisäksi lukeutuu ominaisuus tasolle kaksi tai kolme riippuen toteutuksesta. Tasoilla 0-2 kuljettajan oletetaan olevan jatkuvasti valmiudessa ottamaan ajoneuvon hallinta takaisin ilman erinäistä varoitusta. Tasolla kolme kuljettajan tulee olla valmiudessa ottamaan auton hallinta takaisin järjestelmältä varoituksen jälkeen. Varoitus voi olla esimerkiksi äänimerkki. SAE:n virallisessa määritelmässä esimerkki tällaisesta tilanteesta on onnettomuustilanne maantiellä, minkä johdosta liikennettä ohjaa viranomainen käsimerkein. Kuljettajan tulee siis olla valmiina hallitsemaan autoa lyhyellä varoitusajalla myös kolmannella tasolla (SAE International, 2018).



SAE J3016™ LEVELS OF DRIVING AUTOMATION

	SAE LEVEL 0	SAE LEVEL 1	SAE LEVEL 2	SAE LEVEL 3	SAE LEVEL 4	SAE LEVEL 5
What does the human in the driver's seat have to do?	You are driving whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering			You are not driving when these automated driving features are engaged – even if you are seated in "the driver's seat"		
	You must constantly supervise these support features; you must steer, brake or accelerate as needed to maintain safety			When the feature requests, you must drive	These automated driving features will not require you to take over driving	
What do these features do?	These are driver support features			These are automated driving features		
	These features are limited to providing warnings and momentary assistance	These features provide steering OR brake/acceleration support to the driver	These features provide steering AND brake/acceleration support to the driver	These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met		This feature can drive the vehicle under all conditions
Example Features	<ul style="list-style-type: none"> • automatic emergency braking • blind spot warning • lane departure warning 	<ul style="list-style-type: none"> • lane centering OR • adaptive cruise control 	<ul style="list-style-type: none"> • lane centering AND • adaptive cruise control at the same time 	<ul style="list-style-type: none"> • traffic jam chauffeur 	<ul style="list-style-type: none"> • local driverless taxi • pedals/steering wheel may or may not be installed 	<ul style="list-style-type: none"> • same as level 4, but feature can drive everywhere in all conditions

Kuva 1. Kaavio SAE:n automatisoidun ajamisen eri tasoista. (SAE, 2018)

Kaksi viimeistä, tasot neljä ja viisi kuvaavat autonomista ajoa mikä ei vaadi kuljettajaa tarkkailemaan ympäristöään tai ottamaan ajoneuvoa hallintaansa missään tilanteessa. Esimerkiksi Tampereella Hervannassa toiminut Sohjoa-hankkeen bussi sijoittuisi tasolle neljä, sillä toiminta rajoittuu vain tietylle alueelle ja säälle (Yle, 2016). Teslan Autopilot-laitteiston taso 3.0 lukeutuu tasolle 5 (Tesla, 2020), mutta lainsäädännöllisistä syistä johtuen Autopilot-järjestelmän käyttö edellyttää kuljettajan pitävän otteen ratista. Todellisuudessa kuljettajat eivät aina noudata tätä sääntöä ja internetistä löytyy lukuisia esimerkkejä kuljettajista nukkumassa ratin takana.

Kuva 1 on SAE:n 2018 julkaisema virallinen kaavio automatisoidun ajon eri tasojen ominaisuuksista. SAE:n määritelmää vastaavia erittelyjä on tehty myös Yhdysvaltain kansallisen tieliikenteen turvallisuushallinnon (National Highway Traffic Safety Administration) ja Saksan liikenteen ja digitaalisen infrastruktuurin ministerin toimesta (Luetge, 2017; NHTSA, 2013).

3. Raitiovaunuongelma

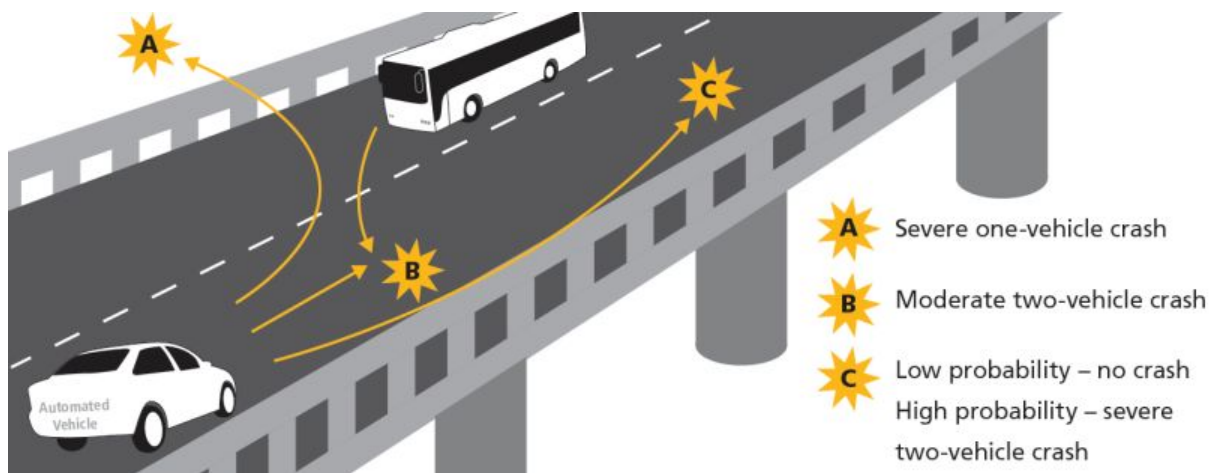
Tässä luvussa käydään läpi mikä on raitiovaunuongelma (*trolley problem*) ja miten se liittyy autonomisiin ajoneuvoihin. Autonomisten ajoneuvojen etiikkaa käsittelevät tutkimukset tuovat lähes poikkeuksetta esiin jonkinlaisen raitiovaunuongelman. Raitiovaunuongelmalla tarkoitetaan tässä asiayhteydessä onnettomuuteen johtavaa tilannetta, missä autonominen ajoneuvo joutuu valitsemaan kahden tai useamman toimintatavan joukosta yhden. Jokainen vaihtoehto aiheuttaa eri määrän vahinkoa. Raitiovaunuongelmien merkityksellisyydestä ei ole yksimielisyyttä. Tässä tutkielmassa viitattujen artikkeleiden kirjoittajista useat puoltavat raitiovaunuongelmien hyötyjä eettisten arvojen määrittelyssä päätöksenteko algoritmeissa (Goodall, 2014a, b; Lin, 2016; Leben, 2016). Toiset pitävät raitiovaunuongelmia merkityksettöminä (Cunneen ja muut, 2020; Nyholm ja Smids, 2018; Himmelreich, 2018).

Keeling (2020) esittää neljä olennaisinta väitettä miksi raitiovaunuongelmia pidetään ongelmallisina. Ensimmäisenä väitteenä raitiovaunuongelmia ei tule tapahtumaan tosielämässä. Vasta-argumenttina tähän Keeling (2020) nostaa esille yleisen kaasulain. Yleinen kaasulaki kuvaa kaasujen käyttäytymistä erilaisissa tiloissa. Vaikka tosielämässä ei ole ideaalikaasuja on yleinen kaasulaki hyödyllinen fysiikassa. Toiseksi raitiovaunuongelmat ja todelliset onnettomuustilanteet eriävät oleellisesti toisistaan, jälkimmäisen sisältäessä epävarmuuksia ja riskejä mitä ei oteta huomioon raitiovaunuongelmissa. Kolmanneksi raitiovaunuongelmat edellyttävät ylhäältä-alas (*top-down*) lähestymistä suunnitteluun. Väitteitä kaksi ja kolme käsitellään luvussa 4. Neljäs väite on, että raitiovaunuongelmat kysyvät väärän kysymyksen. Raitiovaunuongelmat antavat moraalisen vastauksen poliittiseen ongelmaan. Tällä tarkoitetaan, että raitiovaunuongelmien tuottamat eettiset kysymykset vaativat yhteiskunnallista hyväksyntää, mikä on Himmelreichin (2018) mukaan epätodennäköistä ihmisten erilaisten moraalikäsitteiden johdosta. Tässä tutkielmassa lainsäädännöllisiä ongelmia käsitellään Saksan liikenteen ja digitaalisen infrastruktuurin

ministerin julkaiseman autonomisten ajoneuvojen eettisen koodiston ohjeiden kautta (Luetge, 2017). Nämä ohjeet kattavat suuren osan tutkielmassa viitatuissa artikkeleissa esiintyneistä lainsäädäntöön liittyvistä ongelmista.

Raitiovaunuongelmien hyödyllisyys tulee selväksi käsittelemällä yhtä. Otetaan esimerkiksi Goodallin (2014b) demonstroima kuvitteellinen tilanne sillalla. Kuvan 2 vasemmassa alalaidassa oleva autonominen ajoneuvo joutuu tilanteeseen missä vastaan tuleva linja-auto ajautuu yllättäen samalle kaistalle kuin autonominen ajoneuvo. Kolme vaihtoehtoa tilanteessa ovat seuraavat:

- Kaarra vasemmalle ja pois sillalta, mikä aiheuttaa vakavan vamman autonomisessa autossa olevalle henkilölle.
- Hidasta ja törmää vastaan tulevaan bussiin, mikä aiheuttaa kohtalaisia vammoja molemmissa ajoneuvoissa oleville henkilöille.
- Yritä ohittaa bussi oikealta, matala todennäköisyys että bussi korjaa ajolinjansa ja onnettomuudelta vältytään, korkea todennäköisyys että molemmat ajoneuvot kärsivät henkilövahinkoja.



Kuva 2. Goodallin (2014) esittämä raitiovaunuongelma.

4. Päätöksenteko onnettomuustilanteissa

4.1 Seurauseettinen lähestyminen

Tässä luvussa perehdytään toimintasuunnitelmiin jotka pohjautuvat seurauseetiikan alle lukeutuvaan hyötyetiikkaan, selvitetään minkä takia tämä on ohjelmallisesti mahdollista ja mitä ongelmia tällä lähestymistavalla on. Seurauseettisellä lähestymistavalla tarkoitetaan autonomisen ajamisen kontekstissa ajamisen arvon olevan peräisin täysin ajamisen seurauksista. Hyötyetiikka on yksi seurauseetiikan muodoista (Blackburn, 2016). Hyötyetiikka tarkoittaa näkökulmaa siten, että päätökset kuinka autonominen auto liikkuu perustuvat

matemaattisiin laskelmiin hyödyn maksimoimisesta. Autonominen ajoneuvo saattaa esimerkiksi maksimoida ekonomista hyötyä ajamalla siten, että polttoainetta kuluu mahdollisimman vähän ja reitti mitä ajetaan on lyhin.

Kuvan 2 esimerkissä ihmis kuljettaja ei ehtisi tekemään riskianalyysia jonka pohjalta toimia, hän olisi vaistoinsa varassa. Toisin kuin ihminen, autonominen ajoneuvo pystyy edellä mainitussa esimerkissä tunnistamaan linja-auton nopeuden, ennakoimaan todennäköisiä liikeratoja ja muodostamaan toimintasuunnitelman tämän pohjalta. Ajoneuvon reitinsuunnittelu algoritmin tulisi määrittellä mahdolliset tulokset valikoimalle (*range*) reittejä ja asettaa painoarvoja tulosten luotettavuuden perusteella. Luotettavuutta määrittelee muun muassa sensorien keräämä tieto. Algoritmin tavoitteena on turvallisuuden ja liikkuvuuden maksimointi, sekä haittojen minimointi (Goodall, 2014a; Lin, 2016). Kustannusfunktiolla määritellään todennäköisiä tuloksia toiminnan seurauksena. Russell ja Norvig (2003) antavat yhtälön 1 mukaisen esimerkin kustannusfunktiosta:

$$EU(A|E) = \sum_i P(Tulos_i(A) | Tee(A), E) \times U(Tulos_i(A)) \quad (1)$$

Autonomisen ajoneuvon kontekstissa funktion 1 muuttujat ovat:

EU = odotettavissa oleva hyöty (*expected utility*)

E = ajotilanne

U = turvallisuus

A = liikerata

P = liikeradan toteutumisen todennäköisyys, kun toteutetaan liikerata A ajotilanteessa E

Yhtälössä 1 yritetään maksimoida odotettavissa olevaa hyötyä. Liikeradan turvallisuuden laskemisessa käytetään yhtä mahdollista tulosta liikeradan turvallisuudesta. Liikeradan turvallisuuteen vaikuttavat esimerkiksi aineelliset vahingot, fyysiset vammat ja/tai menehtyminen onnettomuuden osallisille. Saksan liikenteen ja digitaalisen infrastruktuurin ministerin julkaiseman autonomisten ajoneuvojen eettisen koodiston toinen ohje määrää seurausetiikkaa noudattavissa järjestelmissä etusijalle asetettavaksi ihmishengen suojelun. Yhdeksännessä ohjeessa kielletään ihmisten arvottaminen ominaisuuksien kuten iän, sukupuolen tai fyysisen olemuksen perusteella (Luetge, 2017).

Näistä rajoituksista nousee esiin kolme ongelmaa. Ensimmäinen on todennäköisyyden laskeminen ihmishengen säilymiselle. Jos kuvitellaan yhden toimintatavoista aiheuttavan törmäyksen jalankulkijaan, ei todennäköisyys kolarista selviämiseen ole välttämättä sama terveellä nuorella aikuisella kuin vanhuksella (Leben, 2017). Toinen ongelma on suojarusteiden huomiointi. Lin (2016) antaa seuraavan esimerkin. Kuvitellaan kaksi ajolinjaa, joista yksi aiheuttaa todennäköisen kolarin kypärää pitävän moottoripyöräilijän kanssa ja toinen ilman kypärää ajavaan moottoripyöräilijään. Jos kypärä otetaan huomioon kustannusfunktiossa, saattaa tämä aiheuttaa suojarusteita pitävälle tienkäyttäjälle epäsuotuisan ratkaisun. Kustannusfunktion määrittely vaatii tietoa muista ajoneuvoista, mikä ei välttämättä ole todennettavissa. Kuvan 1 esimerkissä autonomisen ajoneuvon olisi tiedettävä linja-autossa olevien henkilöiden lukumäärä, että saataisiin todenmukainen kuva

riskeistä. On mahdollista, että julkiset kulkuvälineet, kuten linja-autot ja raitiovaunut ovat tulevaisuudessa yhteydessä autonomisiin ajoneuvoihin, jolloin voidaan välittää tietoa matkustajamääristä, jarrutustehosta ja muusta vastaavasta hyödyllisestä informaatiosta (Goodall, 2014a; Nyholm ja Smids, 2018).

4.2 Velvollisuuseettinen lähestyminen

Tässä aliluvussa tarkastellaan päätöksenteon perustumista velvollisuuseettiseen lähestymistapaan ja käydään läpi mitä hyötyjä ja puutteita tällä lähestymistavalla on. Toisin kuin seurausetiikassa, velvollisuusetiikassa arvoa ei määritellä tietyn tilan saavuttavan tekemisen seurauksena vaan tekojen arvo perustuu teot oikeuttavien sääntöjen noudattamiseen (Blackburn, 2016). Autonomisen ajamisen kontekstissa nämä säännöt voivat olla esimerkiksi tieliikennelakeja. Myös Asimovin lain tapaiset säännöt nousevat esiin yhtenä mahdollisuutena ohjata autonomisen ajoneuvon toimintaa (Goodall, 2014b; Gerdes ja Thornton, 2016). Asimovin laki autonomisten ajoneuvojen asiayhteydessä on yksinkertaisuudessaan seuraava:

1. Autonomisen ajoneuvon ei tule vahingoittaa ihmistä toimillaan tai antaa ihmisten vahingoittua toimetttömyyden takia.
2. Autonomisen ajoneuvon tulee noudattaa sille annettuja määräyksiä silloin kun ne eivät ole ristiriidassa ensimmäisen lain kanssa.
3. Autonomisen ajoneuvon tulee suojella omaa olemassaoloaan niin kauan kun se ei ole ristiriidassa ensimmäisen tai toisen lain kanssa.

Ihmisen lukemana nämä kolme lakia kattavat laajan kirjon mahdollisia vaaratilanteita, mutta tämä vaatii intuitiivista tulkintaa, minkä ohjelmallinen toteuttaminen on erittäin vaikeaa. Pelkästään auton kiihdyttäminen tieliikenteessä asettaa ihmisen vaaraan. Näin abstraktien sääntöjen varassa toimiminen ei ole käytännössä mahdollista. Gerdesin ja Thorntonin (2016) muuttaisivat lakien muotoilua hieman ja lisäisivät siihen neljännen kohdan, mikä mahdollistaisi niiden koneellisen noudattamisen. Tällöin muotoilu olisi:

1. Autonomisen ajoneuvon tulisi välttää törmäämistä jalankulkijoihin ja kevyeen liikenteeseen.
2. Autonomisen ajoneuvon tulisi välttää törmäämistä muihin ajoneuvoihin, paitsi silloin kun törmäämisen välttäminen on ristiriidassa ensimmäisen lain kanssa.
3. Autonomisen ajoneuvon ei tulisi törmätä mihinkään muuhun kohteeseen ympäristössä, paitsi silloin kun törmäämisen välttäminen on ristiriidassa ensimmäisen tai toisen lain kanssa.
4. Autonomisen ajoneuvon tulee noudattaa tieliikennelakia, paitsi silloin kun tieliikennelain noudattaminen on ristiriidassa ensimmäisten kolmen lain kanssa.

Nämä neljä lakia ovat jo huomattavasti kattavammat kuin alkuperäiset kolme ja mahdollistavat turvallisen toiminnan monessa tilanteessa. Ilman neljättä lakia niinkin yleinen tilanne kuin pyöräilijä maantien laidalla aiheuttaisi ongelmia. Ajokaistalta poistuminen sulkuviivan yli pyöräilijää väistäessä ei olisi mahdollista tieliikennelakia ehdottomasti noudattavalle ajoneuvolle, eikä pyöräilijän vauhtia kulkeva ajoneuvo olisi liikenteen sujuvuuden kannalta positiivinen asia.

Käytännössä yksinomaan sääntöihin pohjautuva päätöksenteko ei ole mahdollista. Bauer (2020) ehdottaa kaksitasoista lähestymistapaa autonomisten ajoneuvojen päätöksentekoon. Tällöin autonomisiin ajoneuvoihin voitaisiin ensimmäisellä tasolla sisällyttää paikalliset lait ja normit. Toisella tasolla olisi seurauseettinen, kustannusfunktioihin perustuva päätöksenteko. Ensimmäisen tason säännöt voisivat perustua myös toisen tason toiminnan seurauksiin.

4.3 Nykytilanne ja lähitulevaisuus

Tässä aliluvussa käydään läpi päätöksentekoa onnettomuustilanteissa vuonna 2020 ja lähitulevaisuudessa. Aliluvuissa 4.1 ja 4.2 käsiteltiin tilanteita missä autonominen ajoneuvo havainnoi ja luokittelee ympäristönsä täydellisesti. Tämä ei vastaa todellisuutta, kuten käy ilmi esimerkiksi Arizonassa maaliskuussa 2018 tapahtuneesta onnettomuudesta. Autonomisessa tilassa ajanut Volvo XC90 ajoneuvo ei havainnoinut yöaikaan tietä ylittävää jalankulkijaa, mistä seurasi kuolonkolari. (Kohli ja Chadha, 2018). Cunneen ja muut (2020) pitävät artikkelin kirjoittamisen hetkellä autonomisten ajoneuvojen tärkeimpinä ongelmina havaintokyvyn puutteellisuutta. Heidän mukaansa tulisi keskittyä autonomisten ajoneuvojen havaintokyvyn, luokittelun ja autojen välisen kommunikaation kehittämiseen. Sekä Cunneen ja muut (2020), että Kohli ja Chadha (2018) pitävät konvoluutioneuroverkkoja (*convolutional neural network*) tällä hetkellä nopeimpana ja luotettavimpina menetelminä autonomisten ajoneuvojen kameroiden keräämän tiedon luokitteluun. Aliluvun alussa mainittu Arizonassa tapahtunut kuolonkolari oltaisiin voitu välttää käyttämällä You Only Look Once nimistä konvoluutioneuroverkkoa hyödyntävää kohteen havaitsemisjärjestelmää (Kohli ja Chadha, 2018). Edes luotettavimmat havainnointi ja luokittelujärjestelmät eivät yllä kuin noin 90% tarkkuuteen tällä hetkellä. Tämän ja muiden teknologisten rajoitteiden takia autonomiset ajoneuvot vaativat vielä ihmisen olevan valmiudessa ottamaan ajoneuvon hallinnan takaisin haltuunsa epätavallisissa tilanteissa SAE:n (2018) taksonomian kolmannen luokan mukaisesti (Cunneen ja muut, 2020).

Autonomisten ajoneuvojen toiminnan perustuminen neuroverkkoihin ja muihin koneoppimisen malleihin aiheuttaa itsessään eettisiä ongelmia. Koska suuri osa luokittelun harjoitteluun käytetyistä tietojoukoista kerätään ihmis kuljettajien toiminnasta altistuvat järjestelmät syrjinnälle. Ihmisten on havaittu syrjivän jalankulkijoita muun muassa ulkomuodon, etnisyyden ja iän perusteella. Tämän johdosta kaksi samalla tavalla käyttäytyvää erinäköistä ihmistä voivat altistua erisuurille riskeille jalankulkijana (Cunneen ja muut, 2020; Himmelreich, 2018).

Turvallisuuden parantaminen edellyttää kommunikaation kehittämistä autosta autoon (*vehicle to vehicle, V2V*), autosta infrastruktuuriin (*vehicle to infrastructure, V2I*) ja autosta henkilöön (*vehicle to person, V2P*). Havainnointikyvyn parantaminen täydelliseksi ei ole realistista lähitulevaisuudessa ja kommunikaation avulla on mahdollista täyttää havaintokyvyn puutteet (Cunneen ja muut, 2020; van Loon ja Martens, 2015; Nyholm ja Smids, 2018). Lisääntynyt kommunikaatio herättää kysymyksiä sekä yksityisyydestä että tietoturvallisuudesta. Tietoturvaongelmia on löytynyt esimerkiksi Teslan vikapalkkio-ohjelman kautta toukokuuhun 2020 mennessä 450 kappaletta (Bugcrowd, 2020). Liikenteeseen liittyvän kommunikaation tuottaman suuren datan käsittelyn ennustetaan olevan kriittinen osa liikenneturvallisuutta lähitulevaisuudessa (Maddox ja muut, 2015). Tämä saattaa asettaa älylaitteita karsastavat tienkäyttäjät eriarvoiseen asemaan onnettomuusriskin kannalta. Yksityisyyttä taas ei älylaitteita välttämällä tulevaisuudessa enää saa. Vaikka tienkäyttäjää ei kommunikoi minkään laitteen avulla muihin voidaan häntä seurata ajoneuvojen kameroiden kasvotunnistuksen avulla.

5. Keskustelu

Raitiovaunuongelmat ovat olleet osa autonomisten autojen etiikkaa käsitteleviä artikkeleita ensimmäisistä tutkimuksista aina viimeisiin asti. Osa kokee raitiovaunuongelmien olevan hyödyllinen tapa esittää ristiriidassa olevia epäjohdonmukaisuuksia kun taas toiset kokevat raitiovaunuongelmien olevan liian abstrakteja ja yksinkertaisia tilanteita, mistä ei voida johtaa toimintamalleja algoritmeille onnettomuustilanteissa. Vaikka jälkimmäinen väite pitäisi paikkansa ja autonomisten ajoneuvojen kyvyt ennakoita ja toimia osana liikennettä estävät raitiovaunuongelman kaltaiseen tilanteeseen joutumisen on asiasta keskustelu koettu selvästi tärkeäksi. Esimerkiksi tutkielmassa lainsäädännöllisten ongelmien käsittelyssä käytetyn Saksan eettisen koodiston 20:stä ohjeesta viisi käsittelee raitiovaunuongelmien kaltaisia tilanteita ja Luetge (2017) mainitsee näiden viiden ohjeen kuuluneen kiistanalaisimpiin keskustelunaiheisiin.

Autonomisten ajoneuvojen päätöksentekoprosessia onnettomuuteen johtaneissa tilanteissa käsiteltiin vielä muutama vuosi sitten hyvin teoreettisella ja abstraktilla tasolla. Autonomisia ajoneuvoja ei ollut toiminnassa kuin suljetuilla alueilla eikä onnettomuuksia ollut analysoitavaksi. Tällöin julkaisuissa käsiteltiin pääasiallisesti seurauseettisiä ja velvollisuuseettisiä tai näitä yhdisteleviä lähestymistapoja autonomisten ajoneuvojen päätöksenteon perustana onnettomuustilanteissa. SAE:n (2018) taksonomian kolmannen tason ajoneuvojen yleistyessä kävi ilmi, ettei autonomisten ajoneuvojen kyky havainnoida ja luokitella ympäristöä riitä vielä erottamaan luotettavasti jalankulkijaa puusta (IEEE, 2018).

Koska kustannusfunktioiden laskeminen ja sääntöjen noudattaminen vaativat tarkkaa tietoa ympäristöstä, mitä nykyteknologia ei vielä luotettavasti voi tuottaa, on painopiste alan tutkimuksissa selvästi muuttunut. Sen sijaan että keskityttäisiin pahimpien mahdollisten onnettomuuksien käsittelyyn on keskustelu siirtynyt onnettomuuksien ennakkointiin ja

nykypäivän ja lähitulevaisuuden menetelmien tuottamiin ongelmien käsittelyyn. Kun autonomisten ajoneuvojen lukumäärä kasvaa tulevaisuudessa voi olla, että vakavat onnettomuudet yleistyvät eivätkä onnettomuuteen johtaneet syyt ole niin selviä kuin nykypäivänä ja painopiste muuttuu jälleen. Tähän voi olla vielä pitkä matka. Cunneen ja muut (2020) arvioivat, etteivät julkaisuissa suositut hyötyeettisen lähestymistavan kustannusfunktioita hyödyntävät algoritmit tule olemaan relevantteja lähitulevaisuudessa, sillä teknologian tulee kehittyä vielä valtavasti, että päästään yli havainnoinnin luokittelun ja lokalisoinnin tuottamista ongelmista.

Lähitulevaisuudessa todennäköisempiä kulmakiviä autonomisten ajoneuvojen päätöksenteossa tulevat olemaan V2V, V2I ja V2P kommunikation kehittyminen. Kommunikaation lisääntyessä yksityisyydensuoja ja kyberturvallisuus ovat aiheita joihin kiinnittää huomiota. Tutkielmassa käytetyssä aineistossa molemmista keskusteltiin vain pinnallisesti. Kommunikaatiosta kerättävän tiedon aiheuttamat eettiset ongelmat tulevat vaikuttamaan myös ihmisiin jotka eivät koskaan käytä moottoroituja ajoneuvoja, sillä autonomiset ajoneuvot havainnoivat ympäristöä jatkuvasti.

Lähdeluettelo

- Bauer, W. (2020). Virtuous vs. utilitarian artificial moral agents. *AI & Society*, 35(1), 263–271. <https://doi.org/10.1007/s00146-018-0871-3>
- Blackburn, S. (2016). *The Oxford Dictionary of Philosophy* (3rd edition.). Oxford: Oxford University Press.
- Bugcrowd, (2020). Bugcrowd Tesla. <https://bugcrowd.com/tesla> (Haettu 22.5.2020)
- Cunneen, M., Mullins, M., Murphy, F., Shannon, D., Furxhi, I., & Ryan, C. (2020). Autonomous Vehicles and Avoiding the Trolley (Dilemma): Vehicle Perception, Classification, and the Challenges of Framing Decision Ethics. *Cybernetics and Systems*, 51(1), 59–80. <https://doi.org/10.1080/01969722.2019.1660541>
- DMV. (2020). Report of Traffic Collision Involving an Autonomous Vehicle (OL 316). https://www.dmv.ca.gov/portal/dmv/detail/vr/autonomous/autonomousveh_ol316+ (Haettu 19.3.2020)
- Gerdes, J., & Thornton, S. (2016). Implementable ethics for autonomous vehicles. *Autonomous Driving: Technical, Legal and Social Aspects* (pp. 87–102). https://doi.org/10.1007/978-3-662-48847-8_5
- Goodall, N. (2014a). Ethical Decision Making During Automated Vehicle Crashes. *Transportation Research Record*, 2424(2424), 58–65. <https://doi.org/10.3141/2424-07>
- Goodall, N. (2014b). Machine ethics and automated vehicles. *In Road vehicle automation* (pp. 93-102). Springer, Cham.
- Himmelreich, J. (2018). Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations. *Ethical Theory and Moral Practice*, 21(3), 669–684.

- <https://doi.org/10.1007/s10677-018-9896-4>
- IEEE, (2018). NTSB Investigation Into Deadly Uber Self-Driving Car Crash Reveals Lax Attitude Toward Safety.
<https://spectrum.ieee.org/cars-that-think/transportation/self-driving/ntsb-investigation-into-deadly-uber-selfdriving-car-crash-reveals-lax-attitude-toward-safety> (Haettu 22.5.2020)
- Keeling, G. (2020). Why Trolley Problems Matter for the Ethics of Automated Vehicles. *Science and Engineering Ethics*, 26(1), 293–307.
<https://doi.org/10.1007/s11948-019-00096-1>
- Kohli, P., & Chadha, A. (2018). Enabling Pedestrian Safety using Computer Vision Techniques: A Case Study of the 2018 Uber Inc. Self-driving Car Crash. *arXiv.org*.
https://doi.org/10.1007/978-3-030-12388-8_19
- Leben, D. (2017). A Rawlsian algorithm for autonomous vehicles. *Ethics and Information Technology*, 19(2), 107–115. <https://doi.org/10.1007/s10676-017-9419-3>
- Lin, P. (2016). Why ethics matters for autonomous cars. In *Autonomous driving* (pp. 69-85). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-45854-9_4
- Luetge, C. (2017). The German Ethics Code for Automated and Connected Driving. *Philosophy & Technology*, 30(4), 547–558.
<https://doi.org/10.1007/s13347-017-0284-0>
- Maddox, J., Sweatman, P., & Sayer, J. (2015). Intelligent vehicles + infrastructure to address transportation problems – a strategic approach. In *24th international technical conference on the enhanced safety of vehicles (ESV)*.
- NHTSA. (2013). Preliminary statement of policy concerning automated vehicles. *Washington, DC*, 1-14.
https://www.nhtsa.gov/staticfiles/rulemaking/pdf/Automated_Vehicles_Policy.pdf (Haettu 23.2.2020)
- Nyholm, S., & Smids, J. (2018). Automated cars meet human drivers: responsible human-robot coordination and the ethics of mixed traffic. *Ethics and Information Technology*, 1–10. <https://doi.org/10.1007/s10676-018-9445-9>
- Ramsey, M. (2015). Who’s Responsible When a Driverless Car Crashes? Tesla’s Got an Idea. *Wall Street Journal*, May, 13.
- Russell, S., & Norvig, P. (2003). Artificial intelligence : a modern approach (2nd ed.). *Upper Saddle River (NJ): Prentice Hall*.
- SAE international. (2018). Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles. *SAE International*,(J3016_201806).
https://doi.org/10.4271/J3016_201806
- Tesla. (2020). Autopilot and Full Self-Driving Capability
<https://www.tesla.com/support/autopilot> (Haettu 24.4.2020)
- van Loon, R., & Martens, M. (2015). Automated driving and its effect on the safety ecosystem: how do compatibility issues affect the transition period? *Procedia Manufacturing*, 3, 3280–3285. <https://doi.org/10.1016/j.promfg.2015.07.401>.
- Yle. (2016). Yleisradio uutiset. <https://yle.fi/uutiset/3-9263578> (Haettu 22.2.2020)

Kuvien lähteet

Kuva 1

SAE International, (2018). SAE International Releases Updated Visual Chart for Its “Levels of Driving Automation” Standard for Self-Driving Vehicles.

<https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%E2%80%9Clevels-of-driving-automation%E2%80%9D-standard-for-self-driving-vehicles> (Haettu 18.3.2020)

Kuva 2

Goodall, N. (2014). Ethical Decision Making During Automated Vehicle Crashes. *Transportation Research Record*, 2424(2424), 58–65.

<https://doi.org/10.3141/2424-07>