

## Posudek oponenta bakalářské práce

Název práce: Intention Estimation of Traffic Participants Using Bayesian Network

Autor práce: Antonín Vobecký

Autor posudku: RNDr. Daniel Průša, Ph.D. (Katedra kybernetiky, FEL, ČVUT)

Hlavním cílem posuzované bakalářské práce bylo navrhnout a implementovat Bayesovskou síť pro predikci chování účastníků silničního provozu. K dosažení tohoto byly stanoveny dva dílčí cíle: 1. seznámit se s pravděpodobnostním programováním a Bayesovskými sítěmi, 2. vytvořit simulátor dopravních situací za účelem získání dat pro sestavení a otestování Bayesovských sítí. Práce měla jasnou praktickou motivaci danou evropským projektem H2020 UP-Drive (Automated Urban Parking and Driving).

Všechny body zadání byly splněny. Autor se seznámil s Bayesovskými sítěmi a s pravděpodobnostním programovacím jazykem Figaro. Příslušné poznatky zdokumentoval v první a ve druhé kapitole. Simulátor naimplementoval v programovacím jazyce Java. Pro predikci navrhnul dvě Bayesovské sítě. První, více obecnou, použil s daty získanými ze simulátoru. Druhou síť, zaměřenou na chování chodců v blízkosti přechodu, vytvořil s reálnými daty poskytnutými z projektu UP-Drive. V tomto případě se jedná o výsledek nad rámec zadání, s využitím reálných dat se při zadávání práce nepočítalo.

Práci lze považovat za nadprůměrně rozsáhlou. Rozsáhlá je implementace simulátoru, který podporuje velké množství entit a nabízí bohaté možnosti pro modelování silničního provozu. Obsáhlý je i samotný text práce, který má 66 stran.

Formální i odbornou úroveň textu považuji za velmi dobrou. Vše je napsáno přehledně, s řídkým výskytem gramatických chyb nebo překlepů. Volba zdrojů je přiměřená a vše je korektně citováno. Vyjimkou je pouze kapitola pojednávající o Bayesovských sítích, ve které je četnost chyb vyšší, například:

- str. 4:  $p(X_1, \dots, x_n)$  namísto  $p(X_1, \dots, X_n)$
- str. 5:  $p(x_1, \dots, N)$  namísto  $p(x_1, \dots, x_N)$
- str. 6: chybná nebo nezvyklá notace - suma, v níž proměnná je náhodná veličina  $X_1$
- nejednotnost v označování náhodných veličin (střídavém používání malých a velkých písmen: 'x' a 'X')

Kromě toho je výklad místy veden za použití množství příkladů s absencí formálních definic.

Z hlediska hloubky detailů není moc dobře zpracovaná kapitola o simulátoru. Výklad je příliš zaměřen na implementaci, na úroveň tříd a metod, chybí popis vyšší úrovně, který by lépe vysvětlil zavedené entity a vztahy mezi nimi, bez důrazu na to, že pro implementaci byl použit programovací jazyk Java. Entity a vztahy by pro lepší orientaci bylo dobré zachytit též schematicky.

Zvolený postup pro sestavení Bayesovských sítí považuji ve většině kroků za správný. Dosažené výsledky jsou slibné zejména v případě druhé (jednodušší) sítě naučené na reálných datech. Domnívám se, že na tyto výsledky může být v projektu Up-Drive s úspěchem navázáno.

K návrhu simulátoru a Bayesovských sítí mám následující připomínky (ze kterých zároveň plynou otázky k obhajobě):

- Reakce entity určitého typu (například automobilu) v daném okamžiku je v simulátoru odvozena podle jednotného modelu, který popisuje průměrné chování (str. 25). Není tedy zohledněno, že jednotliví účastníci silničního provozu mohou konstantně projevovat své specifické vlastnosti. Do jaké míry toto ovlivňuje reálnost simulovaných dat?
- V případě, že v blízkosti přechodu je více chodců, je pravděpodobnost, že alespoň jeden z nich má v úmyslu vstoupit na přechod, stanovena jako maximum z pravděpodobností tohoto jevu přes jednotlivé chodce (str. 33). To nezohledňuje skutečnost, že pravděpodobnost roste s počtem chodců. Co bylo důvodem pro výše uvedené zjednodušení a jaký má dopad na kvalitu simulace a predikce?

- První Bayesovská síť (využívající simulovaná data) je dobře popsána pokud jde o její strukturu a odvození parametrů. Není však jasným způsobem prezentováno zhodnocení kvality získaného modelu (tak, jak je to provedeno v případě reálných dat). Popište způsob, jak otestujete spolehlivost predikce na základě simulovaných dat.

Závěrem mohu konstatovat, že předložená práce má po všech stránkách velmi dobrou úroveň, vyjmenované nedostatky jsou převážně menšího významu. Práci doporučuji k obhajobě a navrhuji ji klasifikovat známkou **B – velmi dobře**.

V Praze, dne 12.6.2017

RNDr. Daniel Průša, Ph.D.