

IV127

Soupis provedené práce

jaro 2018

Matěj Vaněk

- analýzy Robomise.cz, téma pokračuje do DP

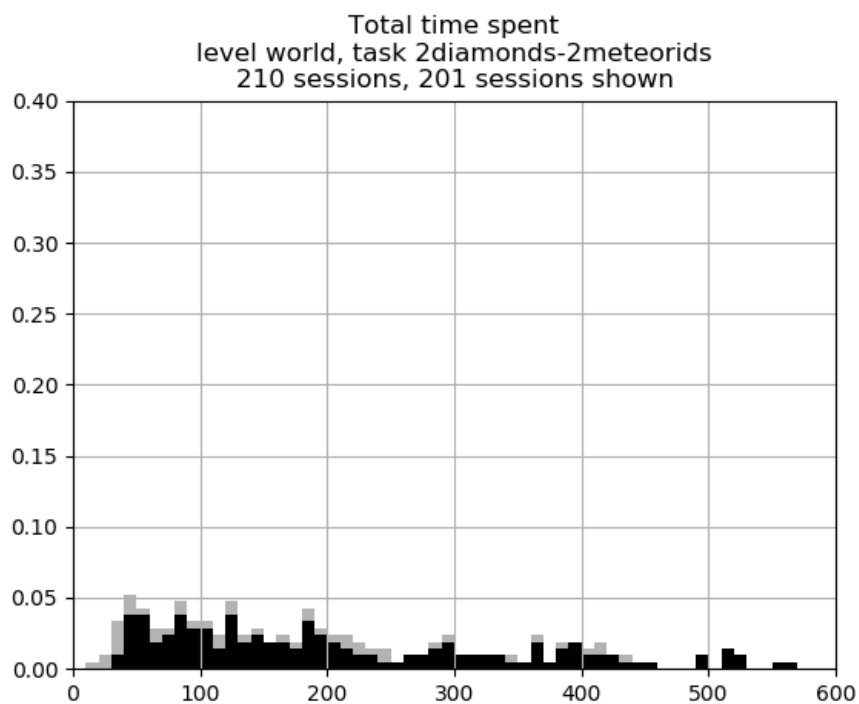
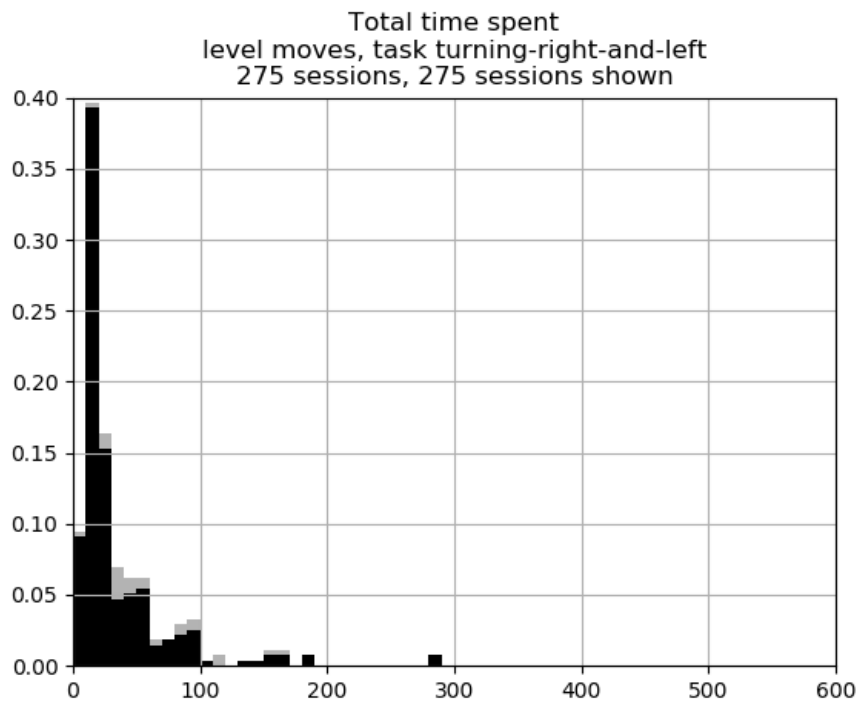
- plán DP:

červen–srpen	osnova práce, hledání a čtení literatury, sestavení množiny zkoumaných veličin a metod jejich měření
srpen–říjen	analyzování metod
listopad	vizualizace veličin, zopakování analýz na konečných datech
prosinec	dokončení, odevzdání

- kód: <https://github.com/matej-vanek/dp>

Histogramy času

- delta mezi operacemi ořezána na 30 s
- černě úspěšné sessions, šedě neúspěšné
- data únor 2018
- většinou lognormální rozdělení, podle očekávání

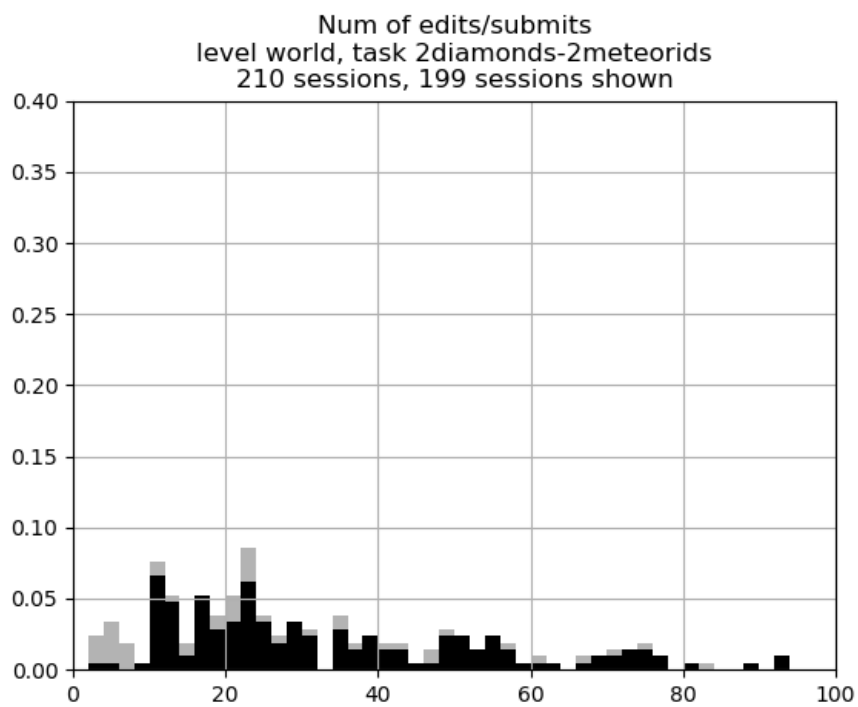
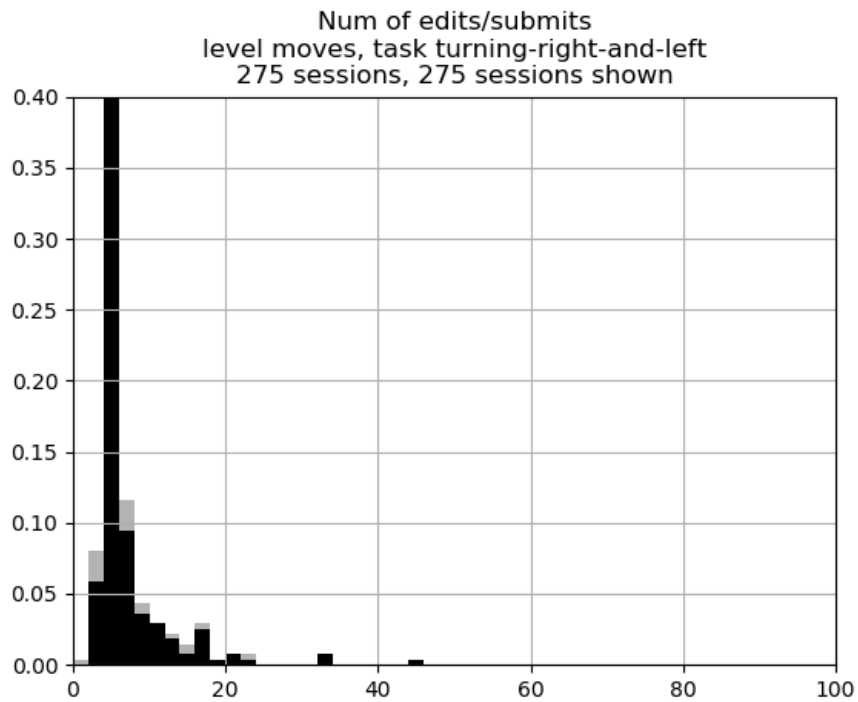


Histogramy operací

- černě úspěšné sessions, šedě neúspěšné

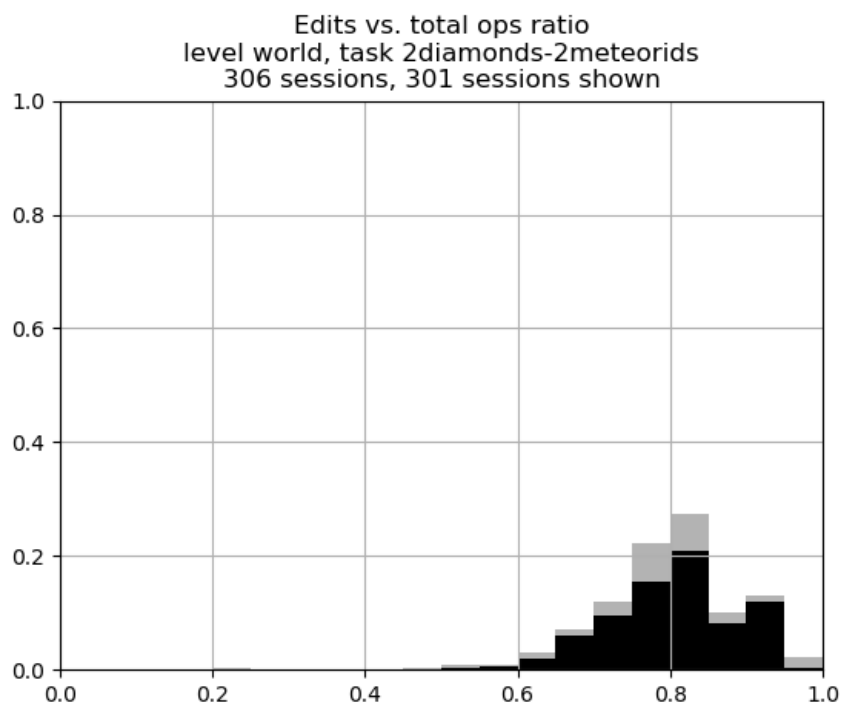
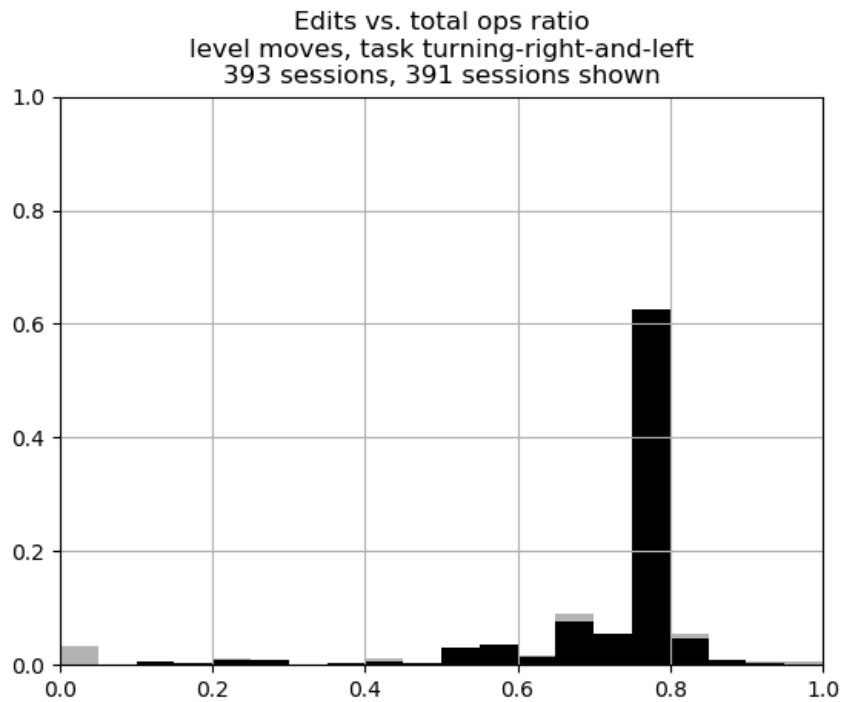
- data únor 2018

- většinou lognormální rozdělení, s větší špičkou než u operací, podle očekávání



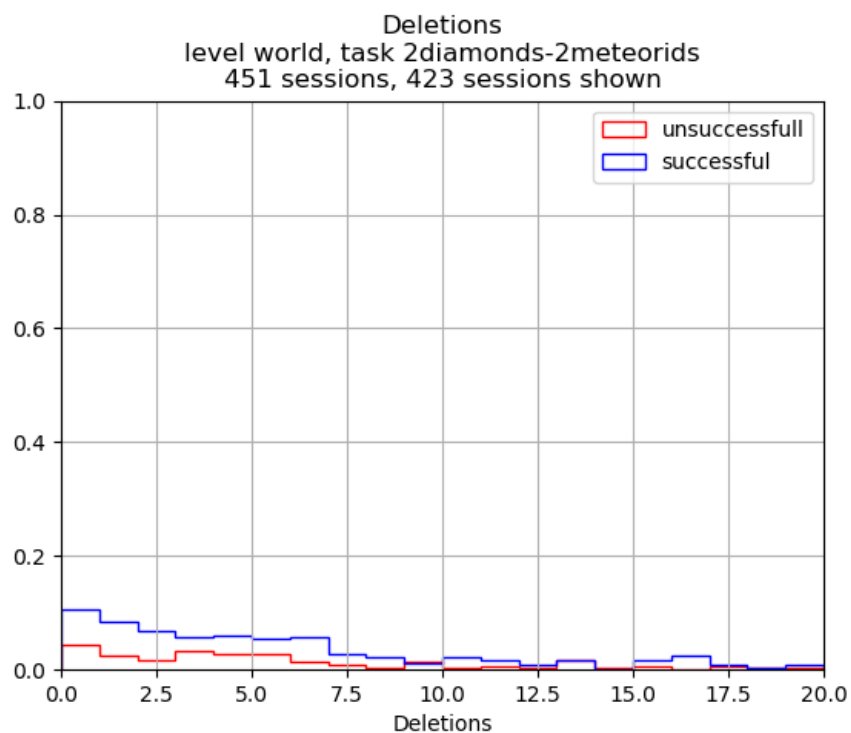
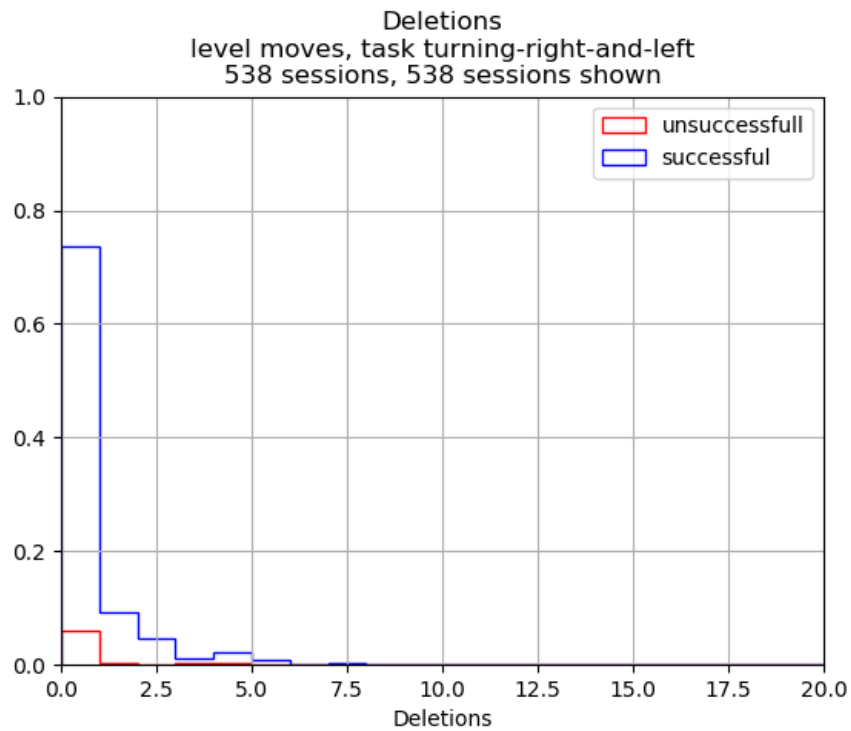
Histogramy poměru edity vs. všechny operace

- černě úspěšné sessions, šedě neúspěšné
- data konec března 2018
- většinou lognormální rozdělení s vrcholem okolo 0.8



Histogramy mazání

- mazání = zkrácení miniRoboCode
- data konec dubna 2018
- kromě lehkých úloh většinou jen pomalu klesající, bez velkých rozdílů úspěšní/neúspěšní



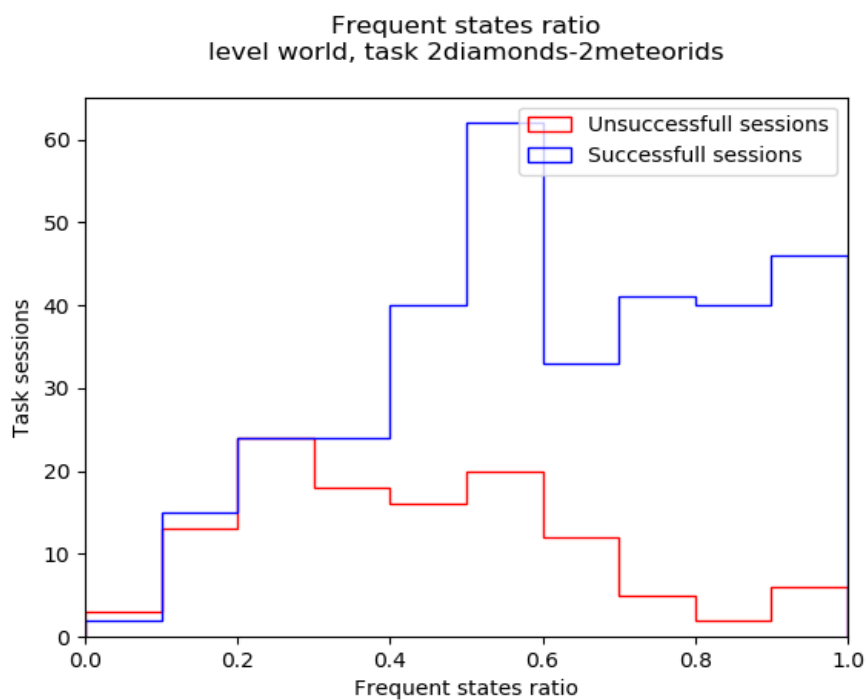
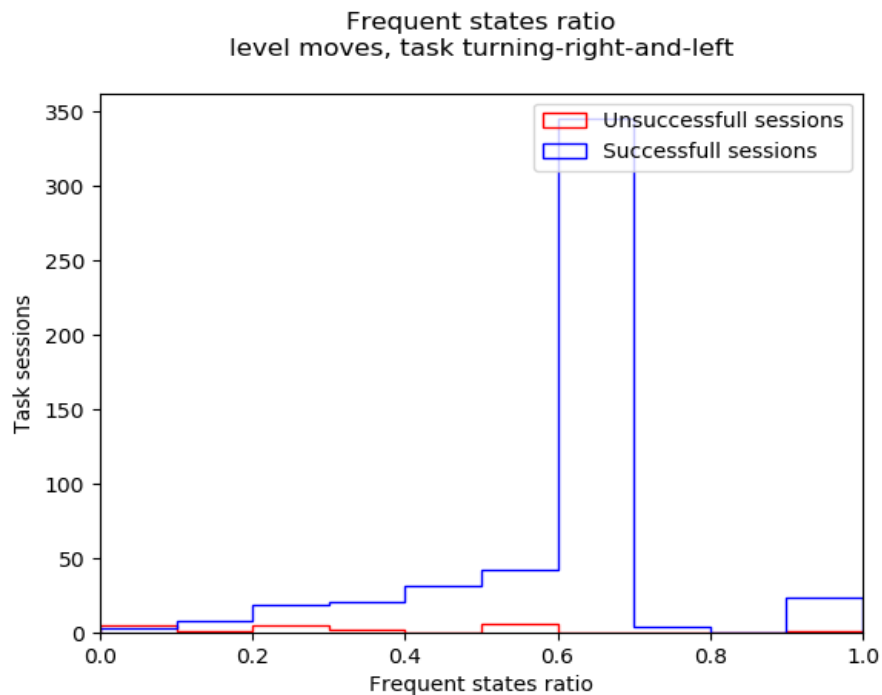
Histogram poměru časté stavy vs. všechny stavy

- časté stavy = ty nejčastěji navštěvované stavy, jejichž návštěvy tvoří kumulativně aspoň 50 % všech návštěv

- data konec dubna 2018

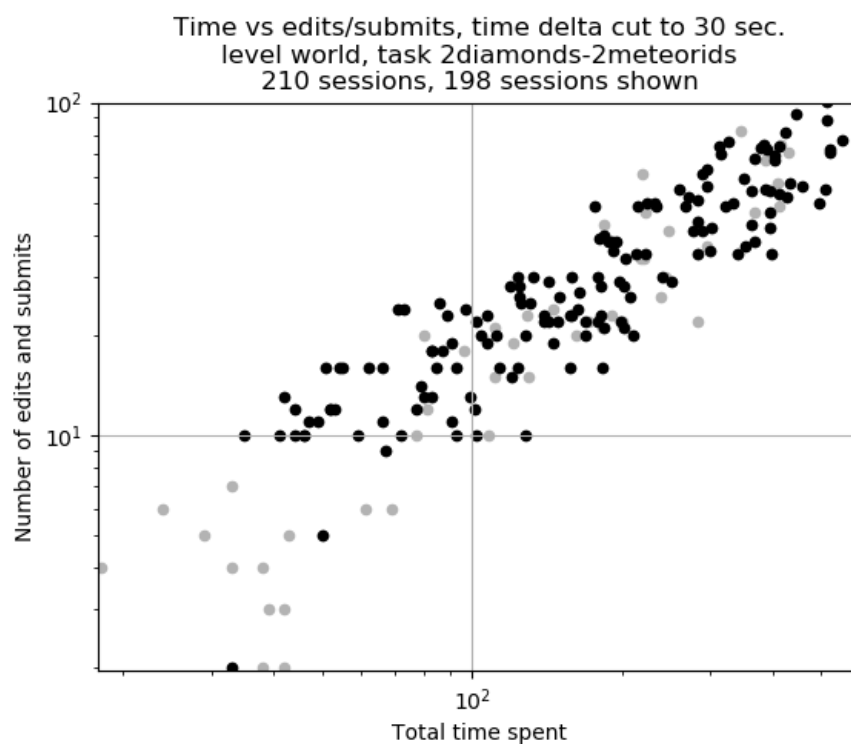
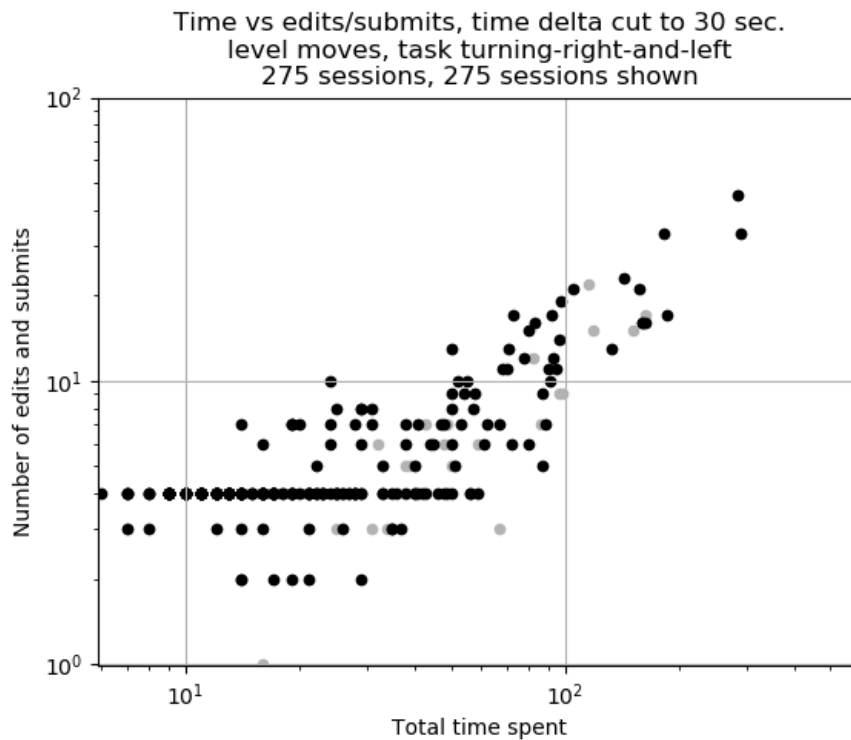
- nepraktická definice častých stavů (změna množiny, časté stavy \neq správné stavy)

- kromě lehkých úloh dost nepravidelné průběhy; pro rozhodování o studentech se moc nehodí, spíš o úlohách



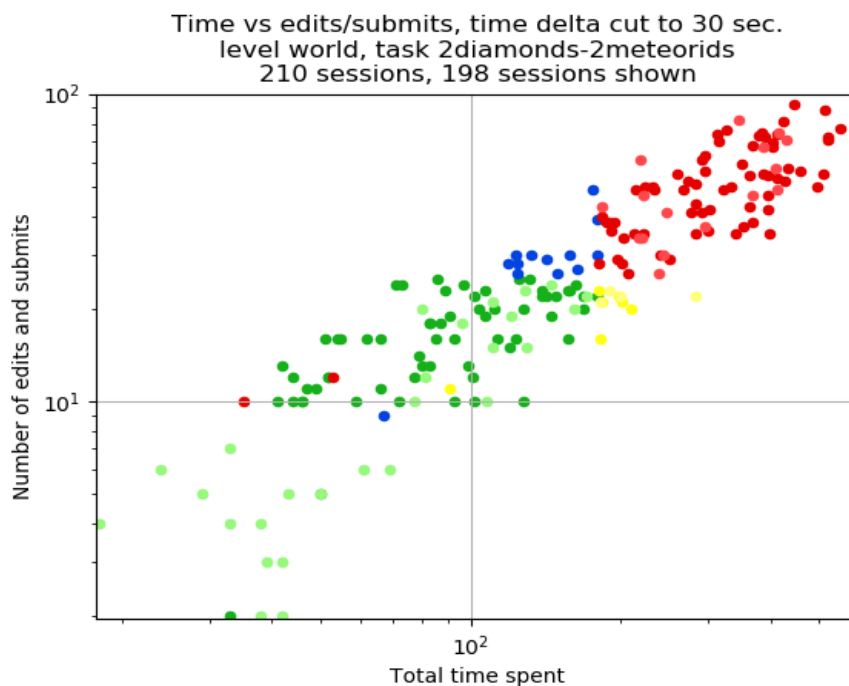
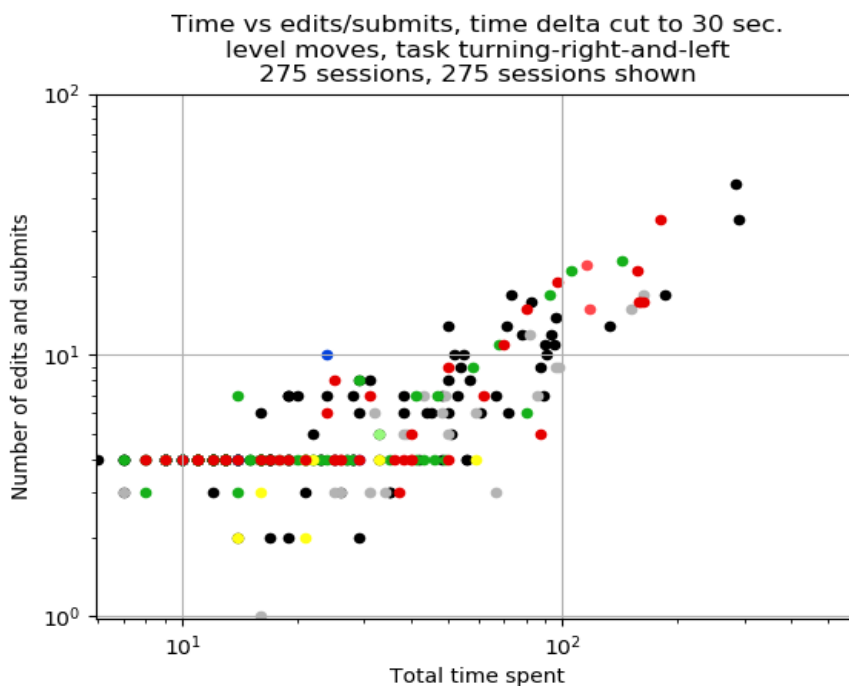
Scatter ploty čas vs. edity

- delta mezi operacemi ořezána na 30 s
- černě úspěšné sessions, šedě neúspěšné
- data únor 2018
- potvrzení, že čas a edity spolu korelují

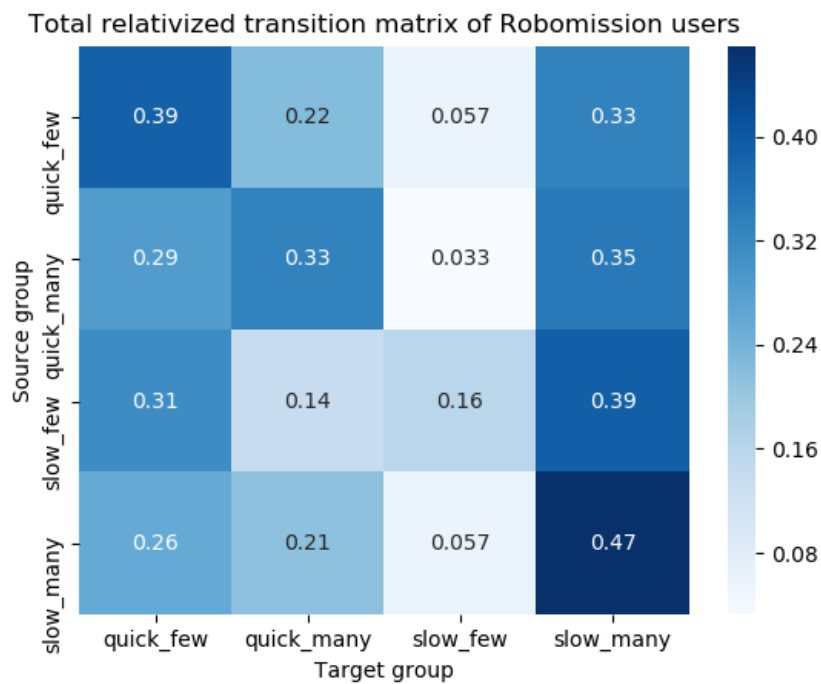


Trajektorie času a operací studentů mezi jednotlivými úlohami

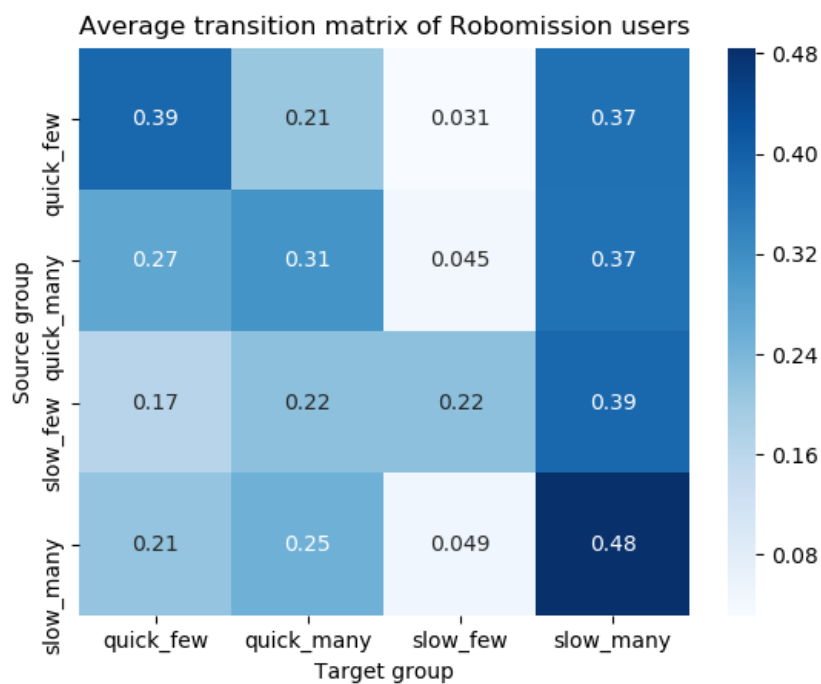
- obarvení určeno podle referenční úlohy
- červeně pomalí – hodně operací, modře rychlí – hodně operací, žlutě pomalí – málo operací, zeleně rychlí – málo operací, černě uživatelé nevyskytující se v referenční úloze
- tmavě úspěšné sessions, světle neúspěšné
- data únor 2018
- nevyváženost skupin, studenti jsou dost nestabilní; hypotéza, že žlutí a modří se drží u „svých“ delších hran



- celková relativní přechodová matice (zrelativizovaná suma všech absolutních matic podle jednotlivých referenčních úloh; nebere v potaz různé množství odpovědí mezi úlohami)



- průměrná relativní přechodová matice (průměr všech relativních matic podle jednotlivých referenčních úloh; bere v potaz různé množství odpovědí mezi úlohami)

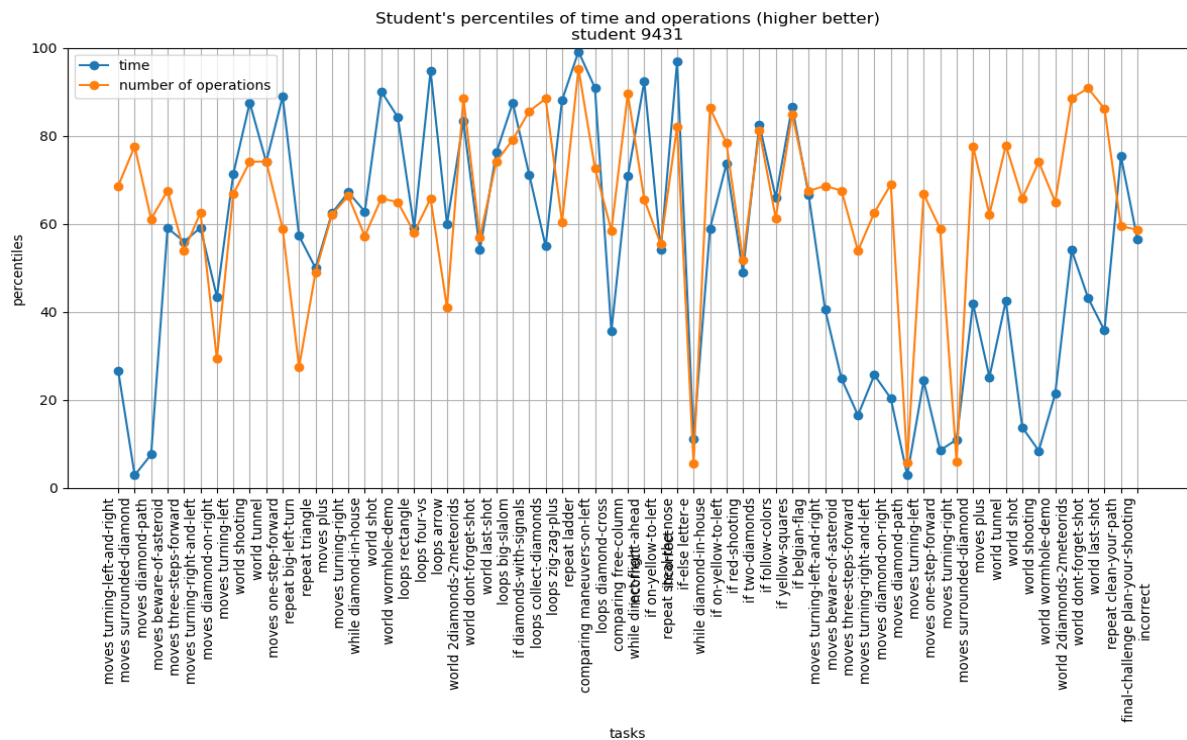
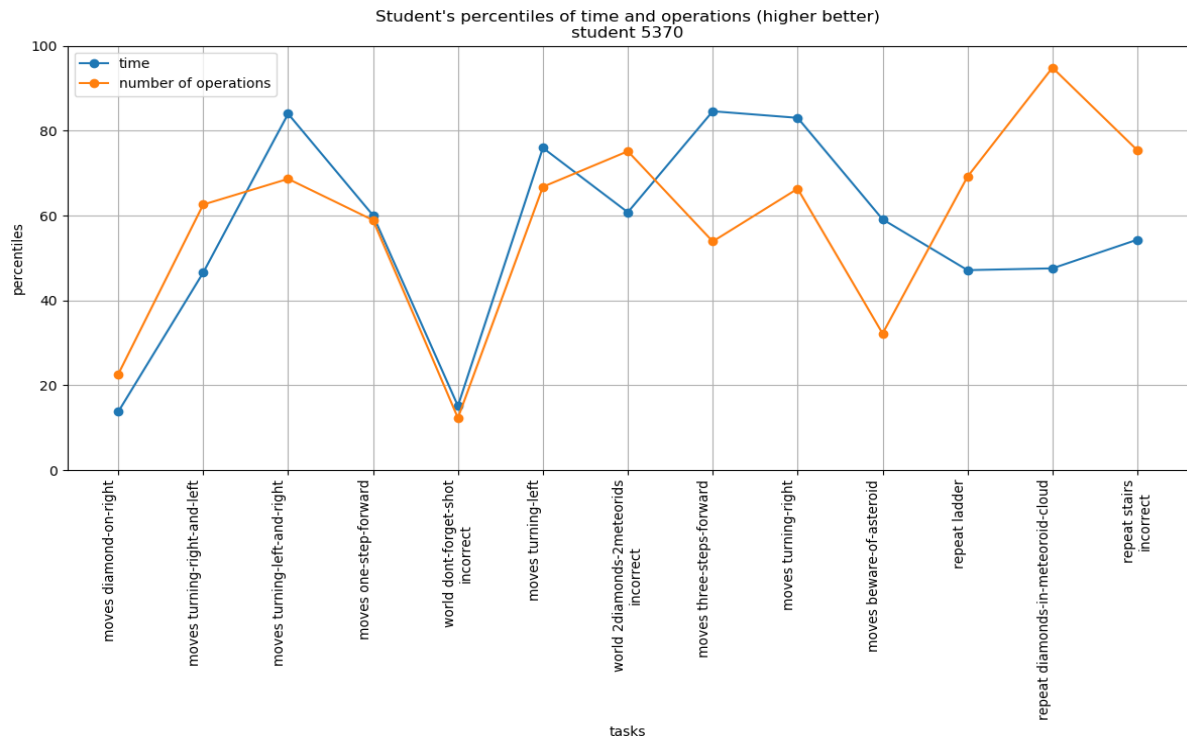


Trajektorie času a operací náhodně vybraných studentů mezi jednotlivými úlohami

- data konec března 2018

- studenti s větším počtem odpovědí sice jsou spíš lepší nebo spíš horší, ale nejsou stabilní

*TODO: totéž, ale jen s úspěšnými sessions



Predikce

- do hodnot 1 (správně, čas pod medián), 0.5 (správně, čas nad medián), případně 0 (špatně)
- metrika RMSE
- výsledky:

metoda	cílová doména	
	[0, 0.5, 1]	[0.5, 1]
náhodně	0.57	0.46
konstantní 0.5	0.39	0.39
konstantní 1	0.51	0.51
poslední čas (>/< medián)	0.48	0.44
poslední operace (>/< medián)	0.50	0.43
poslední čas a operace (>/< medián)	0.49	0.39
poslední čas a operace (prům. percentil)	0.48	0.44
exp. mov. avg. čas (>/< medián)	0.48	0.44
exp. mov. avg. operace (>/< medián)	0.49	0.43
exp. mov. avg. čas a operace (>/< medián)	0.47	0.38
exp. mov. avg. čas a operace (prům. percentil)	0.47	0.43

- závěr: těžké porazit baseline, vyskytují se metodické otázky