

Analýza příbuznosti aproximovaných obvodů

Autor: Kateřina Fořtová

Vedoucí projektu: prof. Ing. Lukáš Sekanina, Ph.D.

Zadání projektu

Vytvořte program, který bude umět pro každou násobičku z zadaného archivu (vč. přesných násobiček, ze kterých CGP vytvořilo aprox. násobičky) vybrat zvolené parametry např. chybové metriky, Area, Power, Delay, Cells uvedené v JSON

Program dopočítá další vlastnosti např. **počet hradel určitého typu, počet podobvodů typu** např. a XOR b XOR c XOR d, (a AND b) XOR (c AND d), **vektory charakterizující dané implementace, zkomprimuje chr soubory, spočítá variabilitu implementace nejvyššího bitu součinu**

Zobrazí závislosti a korelace např. vztah mezi určitou metrikou a počtem hradel XOR, velikostí zkomprimovaného chr souboru, počtem podobvodů určitého typu

Zkuste automatizovaně určit, z jaké přesné násobičky daná aproximovaná násobička vznikla

Implementace

- Nejedná se o evoluční návrh
- Využití **24 918 kusů násobiček** z knihovny EvoApproxLib
- Analýza příslušného JSON souboru s charakteristikou každé násobičky a chromozomu s genotypem násobičky - výsledkem evoluce s využitím CGP
- Implementace v jazyku **Python** s využitím knihovny **Matplotlib** pro vykreslování grafů nebo **PyTorch** pro predikci rodičů aproximovaných násobiček
- **Filtrace JSON souboru** na základě uživatelských preferencí - výběr typu násobičky, počtu filtrovaných objektů, **vybírají se pouze zajímavé metriky užité pro další analýzu**

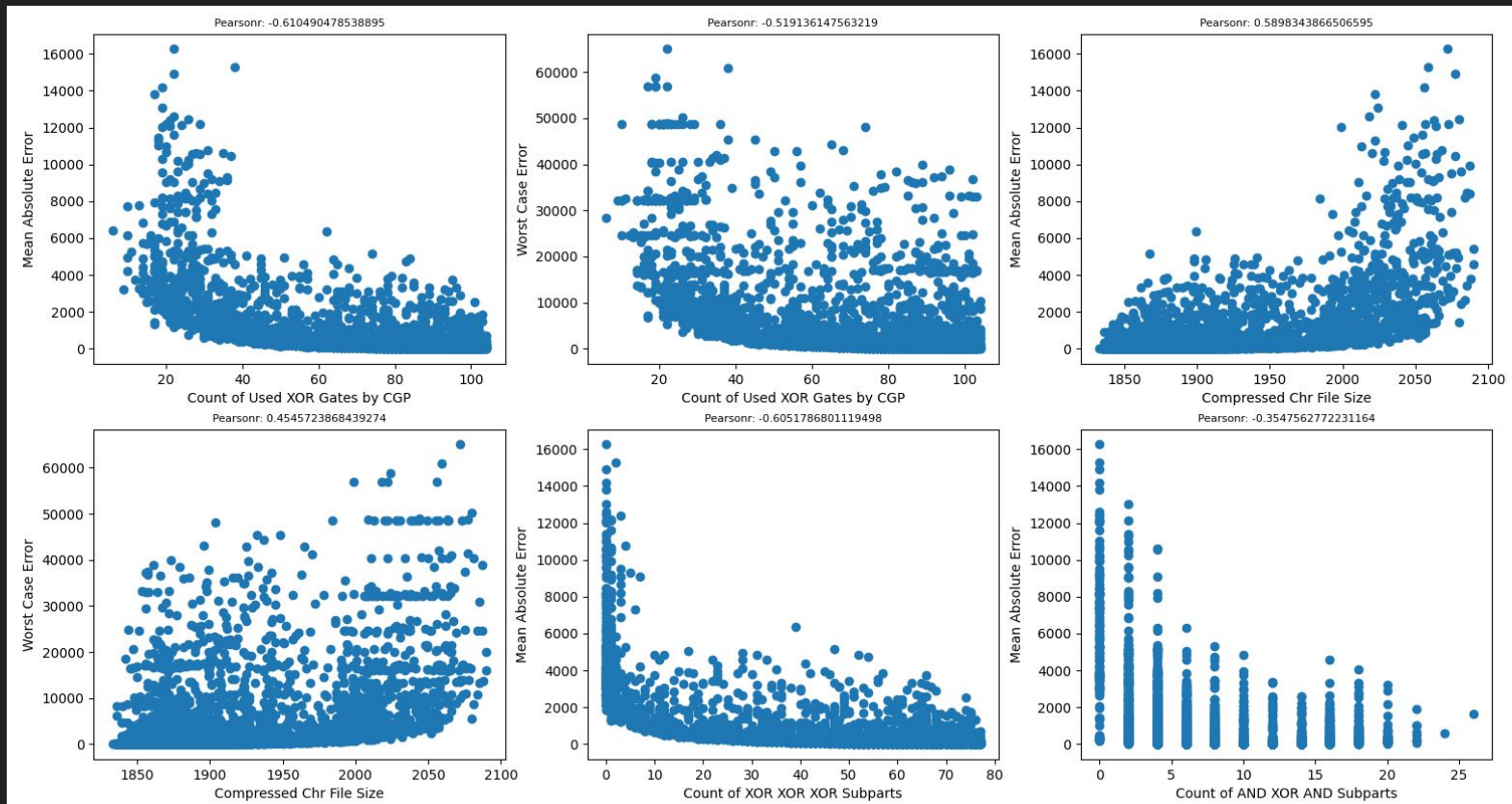
Získání vlastností

Výstupem csv soubor chrFeatures.csv s následujícími informacemi pro každou násobičku z vyfiltrovaného JSON souboru:

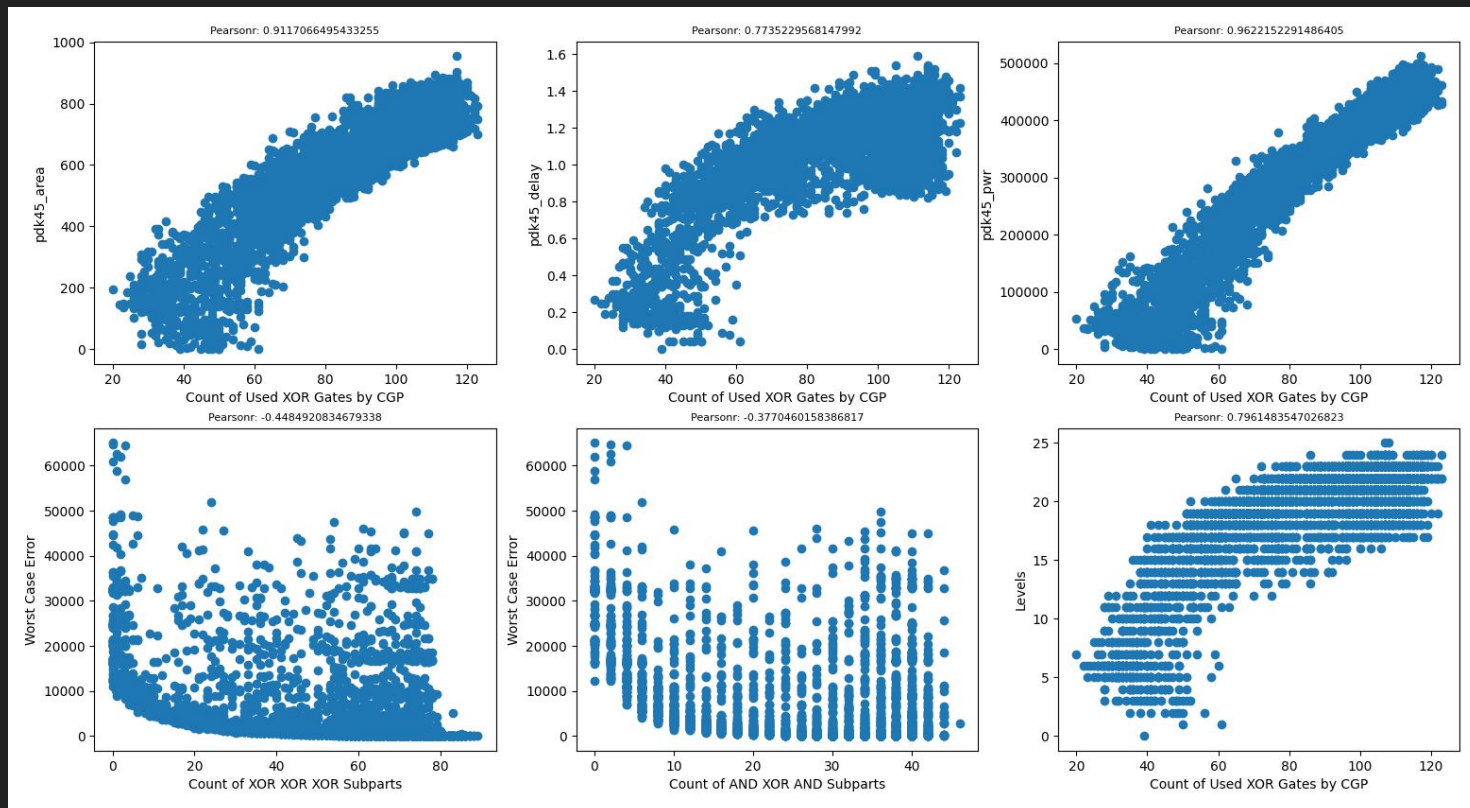
- **Název násobičky**
- **Typy hradel, které využilo CGP** (IDA, INVA, AND2, OR2, XOR2, NAND2, NOR2, XNOR2)
- **Celkový počet hradel, které využilo CGP**
- **Velikost komprimovaného .chr souboru** (chromozomu vytvořeného evolucí)
- **Počet podobvodů** typu a XOR b XOR c XOR d, (a AND b) XOR (c AND d)
- **Variabilita implementace nejvyššího bitu součinu** - přesná násobička počítá nejvyšší bit součinu podle hradel různých typů, tento vektor se získá jako reference a spočítá se vzdálenost od této reference pro jednotlivé aproximativní implementace - u rodičovské násobičky je místo vektoru v tabulce uvedeno "ORIGIN MULTIPLIER"
- **Vektor dvou hodnot** určující, kolik CGP mohlo využít maximálně hradel a kolik skutečně při evoluci použilo

Name	IDA Count	INVA Count	AND2 Count	OR2 Count	XOR2 Count	NAND2 Count	NOR2 Count	XNOR2 Count	Count of used Gates by CGP	Compressed Chr Size	a XOR b XOR c XOR d	(a AND b) XOR (c AND d)	Variability	[All Gates, Used Gates by CGP]
mult8_cgp14ep_ep13107_wc100_wtmcla	26	8	185	66	105	5	6	10	411	2914	73	40	[0, 0, 7, 4, 1, 0, 0, 0]	[490, 411]
mult8_cgp14ep_ep13107_wc100_rcam	0	0	159	47	101	0	0	0	307	1866	73	16	[0, 0, 41, 10, 20, 0, 0, 0]	[320, 307]

Vizualizace grafů závislostí a korelací - aproximované násobičky vzniklé z přesné násobičky typu rcam



Vizualizace grafů závislostí a korelací - aproximované násobičky vzniklé z přesné násobičky typu wtmcla



Rozdíly korelací mezi aproximovanými násobičkami typu rcam a wtmcla

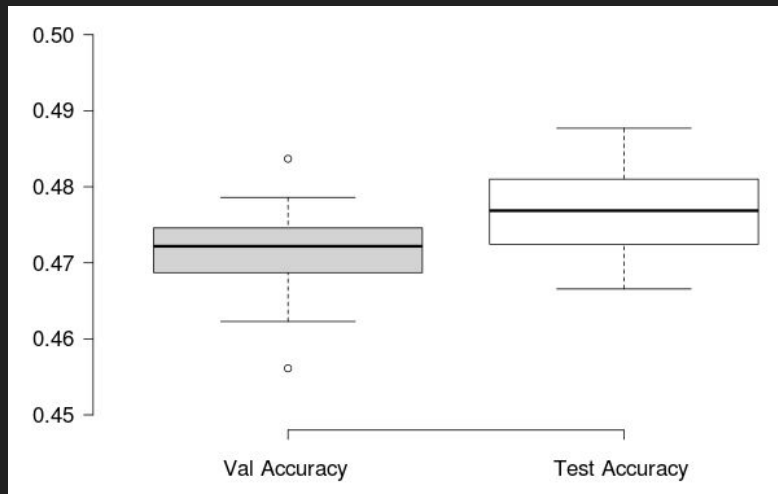
- Výpočet Pearsonova korelačního koeficientu nabývající hodnot v intervalu $\langle -1, 1 \rangle$ - 1 - mezi hodnotami vztah přímé úměry, -1 - mezi hodnotami vztah nepřímé úměry
- Vztah mezi počtem podobvodů ($a \text{ AND } b$) XOR ($c \text{ AND } d$) a Mean Absolute Error - násobičky vzniklé z typu wtmcla mají vyšší vztah pro nepřímou úměru mezi veličinami (-0,61) oproti typu rcam (-0,35)
- Vztah mezi počtem užitých hradel XOR a zpožděním - násobičky vzniklé z typu rcam mají vyšší vztah pro přímou úměru mezi veličinami (0,94) oproti wtmcla (0,77)
- Vztah mezi počtem užitých hradel XOR a počtem log. úrovní od vstupu na výstup - násobičky vzniklé z typu rcam mají vyšší vztah pro přímou úměru mezi veličinami (0,95) oproti wtmcla (0,80)

Predikce rodiče aproximované násobičky

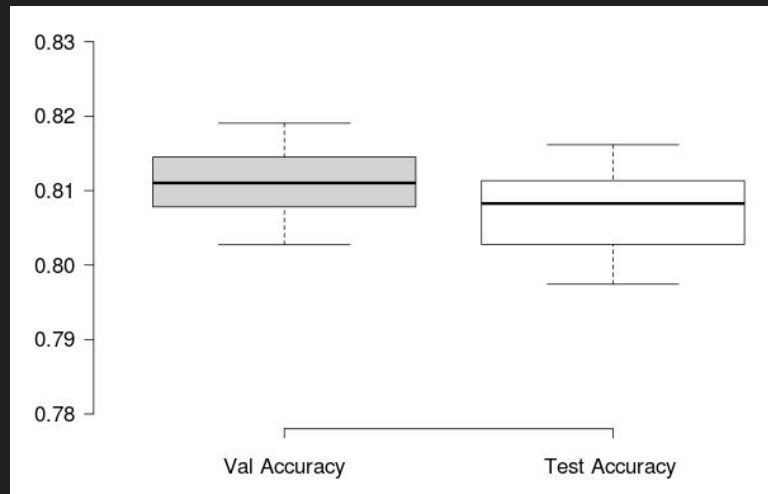
- Zvolení správného typu vektoru rysů
- **MLP klasifikace** využívající aktivačních funkcí ReLu, Softmax a Cross Entropy Loss
- **6 klasifikačních tříd** dle druhu rodičovské násobičky, ze které aproximovaná násobička vznikla:
0 - rcam, 1 - wtm_cla, 2 - csam_rca, 3 - csam_csa, 4 - wtm_csa, 5 - wtm_rca
- **První experimentování** s vektorem 8 hodnot určující počty jednotlivých typů hradel užitých CGP
- v tabulce položky IDA Count, INVA Count, ..., XNOR2 Count - výsledky nejsou dobré
- **Druhé experimentování** s vektorem 8 hodnot určující variabilitu implementace nejvyššího bitu součinu - v tabulce položka Variability - **validační přesnost se zvýšila o zhruba 30 %**
- Náhodný výběr trénovacích, validačních a testovacích dat
- Predikce testovacích dat na modelu s nejlepší validační přesností
- Experimentování s 30 běhy trénování pro každý zvolený vektor
- Trénování po 200 epoch

Predikce rodiče aproximované násobičky - Experimenty s typy zvolených vektorů - Krabicové grafy

Využití vektoru s počty hradel různých typů
využitými CGP:

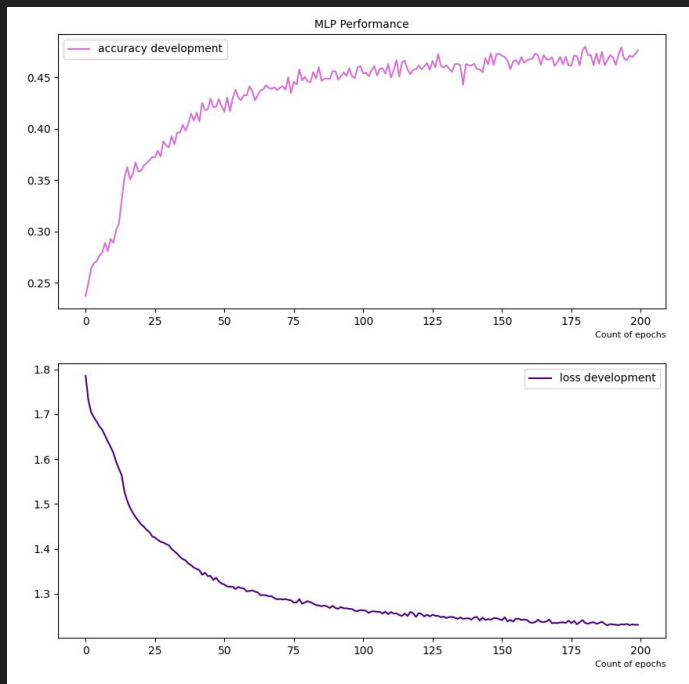


Využití vektoru variability implementace
nejvyššího bitu součinu:

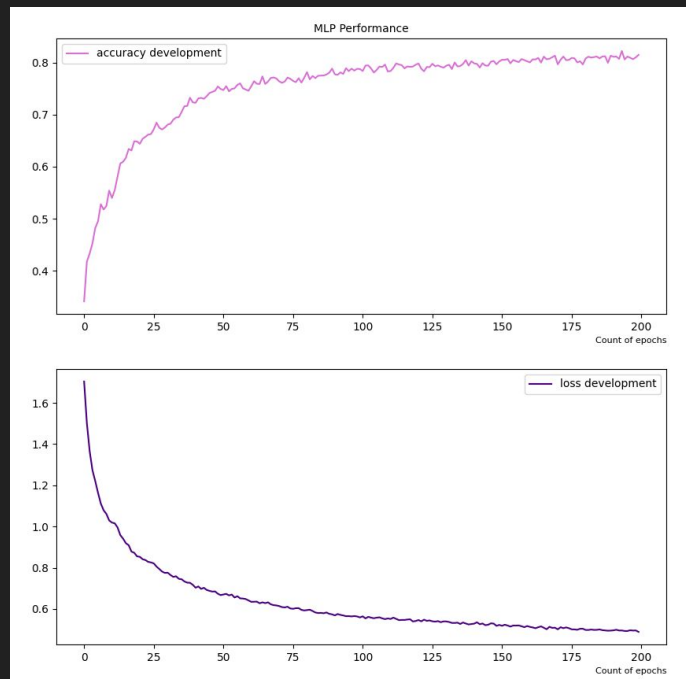


Predikce rodiče aproximované násobičky - Experimenty s typy zvolených vektorů - Grafy validační přesnosti a loss

Využití vektoru s počty hradel různých typů
využitými CGP:



Využití vektoru variability implementace
nejvyššího bitu součinu:



Závěr - zhodnocení

- Naplnění bodů zadání
- Časově náročnější zadání, mnoho podúkolů
- Práce s velkým počtem dat - výpočetně náročnější variabilita implementace nejvyššího bitu součinu, užitá knihovna orjson místo knihovny json pro rychlejší práci s JSON soubory
- Výhodou vyfiltrování určité podmnožiny násobiček - získání výsledků v rozumnějším čase
- Možnost vizualizace mnoha závislostí
- Zakomponování tématu neuronových sítí do predikce rodiče aproximované násobičky