

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Hledání sekvenčních motivů v mRNA selektovaných vazbou na translační iniciační faktory z rodiny eIF4E.
Jméno autora:	Jan Holčák
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra počítačů
Oponent práce:	Doc. Ing. Jiří Kléma, PhD.
Pracoviště oponenta práce:	Katedra počítačů

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Zadání předpokládá komplexní studii stávajících nástrojů pro vyhledávání mRNA sekvenčních motivů a návrh uceleného softwarového řešení pro vyhledávání těchto motivů včetně aplikace na konkrétní data dodaná vedoucím práce. Zadání vyžaduje orientaci v molekulární biologii, bioinformatických databázích i dobré programátorské dovednosti. Zadání proto považuji za náročnější.	
Splnění zadání	splněno s menšími výhradami
<i>Posudte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
Diplomant všechny body zadání splnil. Provedl rešerši, vybral vhodné existující nástroje a integroval je do jednotného prostředí. Na závěr vyhledal sekvenční motivy v úsecích mRNA s vazbou na translační iniciační faktory z rodiny eIF4E. Autor se nezabýval návrhem vlastní metody vyhledání sekvenčních motivů, což zadání považuje za volitelný krok. Soustředil se na integraci řešení stávajících a zde lze již namítnout, že se při této zvolené cestě příliš nezabýval dokumentací celého postupu a ani závěrečná aplikace není příliš hluboká a ilustrativní.	
Zvolený postup řešení	správný
<i>Posudte, zda student zvolil správný postup nebo metody řešení.</i>	
Na základě obecného popisu v textu práce považuji zvolené řešení za správné (zvolené nástroje pro vyhledávání sekvenčních motivů včetně těch relativně nových, reprezentace motivů, práce s kontejnery, API, možnost využití výpočetního clusteru pro práci s rozsáhlými daty). Funkčnost celého řešení jsem ale i vzhledem k chybějícím ilustrativním voláním v omezeném čase nebyl schopen posoudit.	
Odborná úroveň	C - dobře
<i>Posudte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Odborná úroveň práce je dobrá. Oceňuji komplexnost rešerše, orientaci v problému i zvolené softwarové řešení. Na druhou stranu se mi zdá, že výběr nástrojů autor provedl především na základě jejich dostupnosti a hlouběji se nezabýval jejich dalšími parametry. Za velmi stručnou pokládám kapitolu 9, za hlavní výstup lze pokládat tabulku 9.1. Ta není podrobněji popsána, autor v ní bez udání důvodu směřuje E a P hodnoty a nijak nehodnotí nalezené motivy. Chápu, že to může být pro člověka mimo obor složité, ale očekával bych, že se na nalezené motivy bude snažit získat biologickou zpětnou vazbu.	
Formální a jazyková úroveň, rozsah práce	C - dobře
<i>Posudte správnost používání formálních zápisů obsažených v práci. Posudte typografickou a jazykovou stránku.</i>	

Práce má rozumnou strukturu, je psána v češtině bez formálních jazykových problémů. Přesto práce jako celek trpí některými nedostatky: chybí analýza motivů, přesnosti jejich vyhledání a dokumentace xcelého řešení. Autor nijak neřešil všejnou dostupnost svého řešení. Jako celek lze z textu, appendixů a dodaneého kódu jen obtížně odhadnout reálný přínos celé práce.

Výběr zdrojů, korektnost citací

A - výborně

Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posudte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.

Zdroje nejsou zcela v jednotném formátu, to jsem už ale zohlednil při hodnocení formální úrovně práce. V tomto bodě oceňuji širší literatury i její výběr.

Další komentáře a hodnocení

Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.

Práce má integrační charakter, tj. vytváří přehled existujících frameworků a dílčích nástrojů pro vyhledávání sekvenčních motivů a spojuje je do jednotného prostředí. Šlo jistě o časově náročný úkol, technické řešení s použitím Dockeru hodnotím jako správné a uživatelsky přístupné. Samotná instalace byla snadná i pro laika. Přesto bych u práce podobného charakteru čekal několik dalších atributů, které z ní dělají více než jen nástroj pro jedno pracoviště. Mám na mysli zejména zveřejnění celého řešení a s tím spojené rozšíření dokumentace. V minimální podobě by mělo jít o readme soubor, který usnadní použití navrženého pracovního toku. Readme by mohlo obsahovat ilustrační volání pracovního toku, demonstrovat slabé a silné stránky jednotlivých nástrojů apod. Spolu se stručnou Kapitolou 9 pak práce působí dojmem, že se diplomant soustředil především na technické otázky integrace do jednotného prostředí Paraffin a nezabýval se již detailněji možnostmi paralelního použití všech dílčích nástrojů ani jejich dokumentací.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uvedte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Jde o diplomovou práci řešící aktuální a významné téma. Student v ní prokázal schopnost samostatné práce na mezioborovém projektu, ve kterém zcela určitě nebylo snadné se zorientovat. Práce plní všechny body zadání, některé z nich ale podle mého názoru na základní úrovni. Předloženou závěrečnou práci proto hodnotím klasifikačním stupněm .

Otázky: 1. Kolik vlastních řádků kódu jste při řešení napsal? Jsou všechny docker soubory autorské, nebo jsou některé z nich převzaté z dílčích frameworků? 2. Zadání hovoří o návrhu uceleného postupu pro vyhledávání a analýzu sekvenčních motivů v RNASeq NGS datech. Z práce není úplně zřejmé, v čem jde váš návrh za přímé zpřístupnění všech vhodných a dostupných nástrojů do jednoho prostředí. Můžete prosím vysvětlit svůj přístup? Můžete demonstrovat vzorové varianty pracovních toků?

Datum:

Podpis: