Artur Jeż
Instytut Informatyki
Wydział Matematyki i Informatyki
Uniwersytet Wrocławski

28.05.2018 r., Wrocław

Recenzja rozprawy doktorskiej mgra Adama Polaka „Hardness in theory of computing"

1 Ocena

Przedstawiona przez mgra Adama Polaka rozprawa doktorska „Hardness in theory of computing" spełnia wymagania „Ustawy o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki" jak również kryteria zwyczajowe stawiane rozprawie doktorskiej; tym samym wnoszę o dopuszczenie mgra Adama Polaka do dalszych etapów przewodu doktorskiego.

2 Opis i ogólna ocena rozprawy

Przedstawiona rozprawa doktorska składa się z 4 artykułów:

[1] Adam Polak, „Why is it hard to beat $\mathcal{O}(n^2)$ for Longest Common Weakly Increasing Subsequence?", *Information Processing Letters* 132:1–5 (2018).

[2] Lech Duraj, Marvin Künnemann, Adam Polak, „Tight Conditional Lower Bounds for Longest Common Increasing Subsequence", 12th International Symposium on Parameterized and Exact Computation 2017, LIPiCS 89:15:1–15:13 (2017).

[3] Grzegorz Guśpiel, Piotr Micek, Adam Polak, „On an Extremal Problem for Poset Dimension", *Order* (2017).

[4] Grzegorz Gutowski, Konstanty Junosza-Szaniawski, Patryk Mikos, Adam Polak, Joanna Sokół, „Online Coloring of Short Intervals", nie opublikowane.

Zawartość merytoryczna prac, nawet po podzieleniu na współautorów, jest wystarczająca jak na rozprawę doktorską. Prace są połączone wspólnym tematem „trudności w teorii obliczeń". Nadanie takiego tytułu moim zdaniem nie jest właściwe; nie wpływa to jednak na ocenę samych artykułów i rozprawy, dlatego też przedstawię swoje wątpliwości w późniejszej części recenzji, tj. rozdziale 3.4.

Prace [1–2] są ze sobą silnie związane, dotyczą warunkowych ograniczeń dolnych dla problemu obliczania najdłuższego wspólnego (słabo lub silnie) rosnącego podciągu. Wysoko oceniam jakość wyników oraz ich tematykę. Chciałbym zaznaczyć, że dziedzina warunkowych ograniczeń dolnych jest obecnie intensywnie rozwijana i należy pochwalić Adama Polaka, iż samodzielnie dołączył do bardzo konkurencyjnej i szybko zmieniającej się dyscypliny.

Praca [3] dotyczy problemu z dziedziny kombinatoryki ekstremalnej: jak duży musi być poset, aby zawierał podposet rozmiaru n i wymiaru 2 oraz uogólnienie tego wyniku na wymiar $d \geq 3$. Problem ten uważam za naturalny i ciekawy. Doceniam jakość i elegancję podanego dowodu, oraz dobre skorzystanie ze znanych, silnych twierdzeń. Jednak zaliczenie tej pracy do dziedziny informatyki wydaje mi się dość ryzykowne.

Praca [4] dotyczy problemu kolorowania online przedziałów o ograniczonej długości (i grafów przedziałowych): dla danego zbioru przedziałów o długościach z $[1 \mathinner{.\,.} \sigma]$ chcemy pokolorować je online tak, aby każde dwa przecinające się miały różne kolory. Podany jest algorytm (zależący od σ) o asymptotycznym współczynniku konkurencyjności $1 + \sigma$ oraz kilka ograniczeń dolnych na współczynnik asymptotyczny, najlepszy z nich to $5/2$. W tej pracy dużo ciekawsze wydają mi się ograniczenia dolne, warto zaznaczyć, że wkład Adama Polaka był tu dużo większy. Mam pewne zastrzeżenia co do modelu, są one dużo ważniejsze w przypadku algorytmu, dla granic dolnych nie mają większego znaczenia: Z niezrozumiałych dla mnie względów przy liczeniu współczynnika asymptotycznego nie liczymy granicy po wielkości instancji, lecz po liczbie chromatycznej grafu. O ile udało mi się sprawdzić, podana konstrukcja działa też, jeśli będziemy używać tradycyjnej definicji.

Cechą wspólną przedstawionych prac jest konstruowanie różnego rodzaju ograniczeń dolnych. Chciałbym podkreślić, że takie konstrukcje są zwykle trudne i kłopotliwe w dowodzeniu. Na uznanie zasługuje fakt, że Adam Polak zajmuje się głównie tego typu problemami. Mimo tego, że podane konstrukcje są dość skomplikowane, są na tyle starannie przemyślane i zdefiniowane, że dowody dla nich są dużo prostsze, niż by się można spodziewać.

Tematyka prac z rozprawy jest dość rozległa. Z jednej strony świadczy to o dużych zdolnościach doktoranta, jego szerokich zainteresowaniach i oczytaniu jak również o zdolności do stosowania nowych metod i narzędzi (tym bardziej, że Adam Polak jest autorem również innych publikacji, których nie włączył do rozprawy z powodu zbyt dużej rozbieżności tematycznej). Z drugiej pozostawia jednak pewien niedosyt: naturalnym jest oczekiwanie, że doktorant, a w zasadzie każdy naukowiec, po uzyskaniu wyników w jakiejś tematyce zgłębi ją na tyle, iż będzie w stanie uzyskać kilka wyników dotyczących jednej tematyki.

3 Szczegółowe uwagi do poszczególnych artykułów, redakcji itp.

3.1 Prace [1–2]

Prace [1–2] dotyczą warunkowych ograniczeń dolnych dla problemu obliczania najdłuższego wspólnego rosnącego podciągu; w tym wypadku „warunkowy" oznacza, iż danych problemów nie można rozwiązać szybciej (w tym wypadku: istotniej szybciej niż $\mathcal{O}(n^2)$, tj. $\mathcal{O}(n^{2-\epsilon})$) o ile prawdziwa jest hipoteza SETH (Strong Exponential Time Hypothesis), która w uproszczeniu mówi, że problem spełnialności formuł boolowskich nie może być rozwiązany w czasie $\mathcal{O}(2^{\delta n})$ dla $\delta < 1$. Problemy rozważane w przedstawionych pracach znajdują się „pomiędzy" dwoma znanymi problemami najdłuższego wspólnego podciągu (dla którego istnieje bardzo znane kwadratowe warunkowe ograniczenie dolne) oraz najdłuższego podciągu rosnącego (dla jednego ciągu), dla którego znany jest algorytm o czasie działania $\mathcal{O}(n \log n)$. Dla tych problemów znane są algorytmy o kwadratowym czasie działania, jednak ograniczenie dolne dla problemu najdłuższego wspólnego podciągu nie uogólniało się w prosty do rozważanych przypadków. Adam Polak udowodnił, że problem słabo rosnącego wspólnego podciągu nie może być rozwiązany w czasie $\mathcal{O}(n^{2-\delta})$, o ile hipoteza SETH jest prawdziwa. Dowód używa redukcji do problemu Orthogonal Vectors, przy użyciu właściwie dobranych gadżetów. Gadżety są w miarę proste, co oczywiście nie ma sugerować, iż było je łatwo skonstruować.

Razem z Lechem Durajem oraz Marvinem Künnemannem rozszerzyli ten wynik również do przypadku ściśle rosnących podciągów oraz kilka powiązanych problemów. Dowód używa dużo bardziej skomplikowanych konstrukcji kombinatorycznych. Wymóg, aby podciągi były ściśle rosnące nastręcza dużo problemów technicznych. Zaproponowane rekurencyjnie zdefiniowane separatory są ładnym i eleganckim rozwiązaniem dla tych trudności.

Nie mam żadnych uwag do tych prac: są wyczerpujące, dobrze i przejrzyście napisane, podobają mi się techniki i uzyskane wyniki, które dobrze wpisują się w najnowsze trendy w informatyce; w szczególności konkurencja w tej tematyce jest duża. Jeszcze raz chciałbym pochwalić Adama Polaka za wybór tematyki i uzyskane wyniki.

3.2 Praca [3]

Wymiar posetu to najmniejsza liczba naturalna d, taka że dany poset można zanurzyć w $\mathbb{R}^d$; dla danej liczby n szukamy $f(n)$: najmniejszej takiej liczby, że każdy poset rozmiaru $f(n)$ zawiera poset wymiaru 2. Łatwo pokazać, że $\sqrt{n} \leq f(n) \leq n$, w pracy pokazane jest ograniczenie górne $f(n) \leq 4n^{2/3} + o(n^{2/3})$ oraz uogólnienie tego twierdzenia: $f_d(n) = \mathcal{O}\left(n^{\frac{d}{d+1}}\right)$ (gdzie f_d jest analogiczną funkcja przy rozważaniu zawierania posetu wymiaru d). Sami autorzy wskazują, iż ich intuicja mówi, że najgorszym przypadkiem (dla wymiaru d) jest krata d+1 wymiarowa. Sama konstrukcja posetu jest zbliżona do kraty, jednak nią nie jest (co być może należało explicite zaznaczyć w samej pracy). Konstrukcja dla przypadku $d = 2$ jest elementarna, (co znowu nie implikuje, że prosta); dla ogólnego przypadku d w dowodzie wykorzystane jest twierdzenie Klazara-Marcusa (choć o ile dobrze rozumiem deklarację współautorstwa, autorzy mieli też dowód nie używający tego twierdzenia).

Ogólnie podana konstrukcja mi się podoba, tak samo jak i sam problem. Moim głównym zastrzeżeniem jest nikły związek samej pracy z (nawet bardzo szeroko pojętą) informatyką.

3.3 Praca [4]

Punktem wyjściowym jest problem kolorowania grafów, w którym ograniczamy się do grafów przedziałowych (tzn. takich, dla których istnieje zbiór przedziałów, takich że wierzchołki odpowiadają przedziałom, a krawędzie przecinaniu się przedziałów). Następnie ograniczamy się do grafów, które można reprezentować jako zbiór przedziałów o długościach ze zbioru $[1\,..\,\sigma]$, dla ustalonego σ. W rozważanym problemie algorytm otrzymuje reprezentację grafu w postaci zbioru odcinków i musi działać online, tzn. po otrzymaniu każdego przedziału musi pokolorować go i nie może potem tego koloru zmienić. Algorytm oceniamy pod względem współczynnika konkurencyjności, tj. porównujemy wynik algorytmu z najlepszym możliwym. Algorytmy online są ugruntowaną dziedziną algorytmiki, która dobrze modeluje problemy, w których główną trudnością jest nieznajomość przyszłości.

Problem badany w pracy był wcześniej badany dla grafów przedziałowych, jednak bez ograniczeń na długość przedziałów w reprezentacji. Praca przedstawia z jednej strony algorytm $1 + \sigma$ konkurencyjny, a z drugiej podaje ograniczenie dolne $5/2$.

Mam kilka poważniejszych uwag, choć nie zarzutów, dotyczących tej pracy. Po pierwsze: szczegóły modelu: przy określaniu asymptotycznego współczynnika ważne jest, co uznamy za rozmiar instancji. W pracy z niezrozumiałych dla mnie przyczyn jest nią nie ilość przedziałów, lecz liczba chromatyczna grafu. Dlaczego? To potencjalnie poważnie zmienia model. Być może istnieje jakiś powód, lecz wypadałoby go podać. Po drugie, algorytm zakłada, że zna wartość σ (co jest akceptowalne) oraz dostaje zbiór przedziałów, a nie sam graf przedziałowy. To czyni problem prostszym, bo reprezentacja przy użyciu przedziałów nie jest jedyna. Ponadto, sam algorytm zależy od parametru b, który w optymalnym przypadku zależy od reprezentacji σ jako ułamka. To sprawia, że działanie algorytmu bardzo zmienia się wraz z drobnymi zmianami σ, co jest raczej niepożądane. (Jako że wkład Adama Polaka dotyczy głównie ograniczeń dolnych, zarzut ten nie dotyczy właściwie jego wkładu w pracę.) Po trzecie, część dowodów dotycząca granic dolnych jest po prostu źle napisana: konstrukcja i pojawiające się liczby są poprawne, ale same argumenty zdają się dotyczyć innej wersji konstrukcji (da się na szczęście odtworzyć właściwe).

Co do samych ograniczeń dolnych, opierają się one z jednej strony na wcześniej znanym ograniczeniu oraz na strategiach przekształcania ograniczeń w lepsze, co samo z siebie jest ciekawym i eleganckim pomysłem. Kolejne, coraz lepsze ograniczenia uzyskiwane są przez kolejne ulepszenia tej strategii (separacja dwustronna zamiast jednostronnej, wielokrotne powtarzanie, itp.) oraz poprawioną analizę (podział możliwych działań algorytmu na parę przypadków). W ostatecznej wersji uzyskane jest ograniczenie 5/2, które jest wkładem Adama Polaka w tę pracę.

Jakkolwiek sam algorytm budzi we mnie umiarkowany entuzjazm, ograniczenia dolne skonstruowane są przy użyciu ciekawego i kombinatorycznie nietrywialnego pomysłu i zasługują na uznanie.

3.4 Uwagi redakcyjne itp.

Na sam koniec chciałbym przekazać trochę uwag redakcyjnych. Jak już wspomniałem, nie wpływają one na merytoryczną ocenę pracy.

Mam pewne wątpliwości, co do wybranego tytułu rozprawy doktorskiej: prace [1–2] należą do głównego nurtu obecnych badań w dziale złożoności obliczeniowej, trudno jest jednak dopatrzeć się związków pracy [3] z teorią obliczeń. Kwestia ostatniej pracy [4] jest mniej oczywista, jednak tradycyjnie teoria obliczeń zajmuje się raczej wyrażalnością przez różnego rodzaju formalizmy czy też zasobami, które do wyrażania są potrzebne. Model obliczeń online i analizy kompetytywnej słabo współgra z tak pojętą teorią obliczeń.

Prace [1–3] zawierają drobne błędy językowe i mieszczą się pod tym względem w normie. Niestety, fragmenty pracy [4] wymagają przeredagowania. Zapewne ma to związek z tym, iż praca ta nie jest jeszcze opublikowana i nie przeszła przez proces recenzji, wydaje się jednak, że obecne tam błędy można wychwycić nawet przy dość pobieżnym czytaniu pracy i w związku z tym powinny były one być naprawione przed złożeniem rozprawy doktorskiej.

Poza samymi artykułami rozprawa zawiera też dłuższy tekst prezentująca zawartość rozprawy w krótszej i bardziej przystępnej formie. Tekst ten w większości składa się w wielu częściach z niemal dosłownych cytatów ze wstępów z artykułów będących rozprawą. Nie uważam tego za właściwe: rozprawa doktorska daje możliwość namysłu i przedstawienia uzyskanych wyników z szerszej perspektywy. Ta możliwość została w dużej mierze stracona.

Poza wspomnianymi już fragmentami istniejących wstępów doktorant przedstawił również ogólne wprowadzenie do dziedzin, których dotyczą pracę. W przypadku złożoności obliczeniowej, zaprezentowany punkt widzenia wydaje mi się kontrowersyjny: główną rolą ograniczeń dolnych jest ma być wykazanie, że nie ma potrzeby dalej ulepszać algorytmów osiągających optymalne złożoności. Wydaje się, że jednak głównym celem badań w dziale złożoności obliczeniowej jest klasyfikacja (problemów, typów zasobów, typów urządzeń czy formalizmów) a nie wskazywanie, że nie ma już co poprawiać algorytmów dla danego problemu. Twierdzenie, że hierarchia problemów wynikająca z badań dla warunkowej złożoności obliczeniowej jest ciekawsza czy trudniejsza niż hierarchia dla problemów między P i NP też jest kontrowersyjne. Pada też stwierdzenie, że jednym z niewielu bezwarunkowych ograniczeń dolnych jest ograniczenie $\Omega(n \log n)$ dla algorytmów sortujących (w modelu drzew decyzyjnych); po pierwsze, zachodzi ono nawet w silniejszych modelach (drzew arytmetycznych a nawet algebraicznych) po drugie, ciężko się zgodzić z opinią, że takich ograniczeń jest niewiele.

4 Podsumowanie

Pracę ogólnie oceniam pozytywnie, drobne problemy redakcyjne nie wpływają na ocenę. Doceniam, iż Adam Polak zajmuje się ograniczeniami dolnymi, które są zwykle trudne w konstrukcji i dowodzeniu, oraz że jest w stanie samodzielnie prowadzić badania w aktywnie rozwijającej się dziedzinie informatyki. Trochę szkoda, iż nie kontynuował on swoich prac w którymś kierunku, lecz zawsze zmieniał zainteresowania badawcze. Nie mam wątpliwości, że rozprawa jest bardzo dobrą podstawą do nadania tytułu doktora nauk matematycznych w dziedzinie informatyki.

Artur Jeż