Algorytmy poprawy czytelności formalnych rozumowań zapisanych w systemie naturalnej dedukcji

Autoreferat rozprawy doktorskiej

Karol Pąk
Instytut Informatyki,
Uniwersytet w Białymstoku

1 Wprowadzenie

Matematyka jest dziedziną wiedzy, która wnosi do przestrzeni społecznej i kulturowej nie tylko konkretną wiedzę, ale też i poczucie, iż jej wyniki są wolne od błędów. Współczesna cywilizacja podejmuje coraz więcej wyzwań, w których standardy działań nie dopuszczają błędów. Dotyczy to szczególnie oprogramowania działającego we wrażliwych strefach życia ludzkiego (np. oprogramowania elektrowni jądrowych, oprogramowania sterującego samolotem itp.). Wielkość takiego oprogramowania sprawia, iż zapewnienie jego poprawności bez wsparcia komputerów jest nierealne. Włączenie w ten proces osiągnięć matematyki tak, aby nie były one najsłabszym ogniwem, wymaga sformalizowania ich w precyzyjnej, akceptowalnej przez komputer postaci.

Wykorzystanie komputerowej weryfikacji w naukach matematycznych jest dobrze znanym sposobem drastycznego ograniczenia lub eliminacji błędów w formułowanych rozumowaniach. Konieczna do użycia komputera szczegółowość wszystkich kroków oraz możliwych przypadków, uniemożliwia co prawda popełnienie drobnych błędów, ale równocześnie zwiększa długość rozumowania. Problem ten istnieje od czasu pierwszych prób formalizacji matematyki, przyczyniając się do zaniechania wielu z nich. Spowodował on m.in. porzucenie dalszych prac nad kolejnym tomem dzieła *„Principia mathematica"* [22] tworzonego przy ekstremalnie rygorystycznym podejściu do formułowaniu definicji i dowodów, reprezentowanego przez A. N. Whiteheada i B. Russella. Również doprowadził on do odwrócenia się świata matematyki od ambitnego projektu grupy naukowców występujących pod pseudonimem pseudonimem Nicolas Bourbaki.

Zatem począwszy od pierwszych prób formalizacji dowodów matematycznych, procesowi temu zawsze towarzyszył problem nieczytelności struktury tak uszczegółowionych rozumowań. Zdaniem niektórych koncepcje systemów formalnych umożliwiające jasne strukturalizowanie dowodów, zapisanych w systemie dedukcji naturalnej, przedstawione przez G. Gentzena [7] lub S. Jaśkowskiego [10] utrudniają, czy wręcz uniemożliwiają, zastosowanie ich do wiernej reprezentacji dowodów twierdzeń prowadzonych w praktyce przez matematyków [17].

Rozkwit formalizacji był możliwy dopiero przy wykorzystaniu komputerów do weryfikacji poprawności tak uszczegółowionych rozumowań. Idea komputerowo wspomaganej automatyzacji i weryfikacji formalnych rozumowań matematycznych stanowiła bowiem punkt wyjścia do zbudowania systemów prowadzenia rozumowań formalnych, tj. Coq, Isabelle/Izar, HOL Light, Mizar. Wykorzystanie komputerów nie tylko zlikwidowało problem sprawdzania poprawności, ale umożliwiło również prowadzenie rozważań formalnych na bardziej ogólnym poziomie, który jest bliższy temu, jaki występuje w nieformalnych dowodach matematycznych. Pomimo zbudowania takich systemów, zapisane w nich formalne rozumowania są nadal nieczytelne dla przeciętnego

użytkownika. Implikuje to konieczność dalszych badań w zakresie metod poprawy czytelności. Zagadnieniom tym została poświęcona prezentowana rozprawa doktorska.

Zasadniczą przyczyną tego stanu rzeczy jest większa niż w normalnej praktyce matematycznej liczba kroków, która jest jednak konieczna do zapewnienia poprawności wywodów formalnych (zależność między długością dowodu nieformalnego i formalnego charakteryzowana jest tzw. współczynnikiem de Bruijna [3, 24]).

Analiza doświadczeń związanych z rozwojem i konserwacją bazy Mizar Mathematical Library (MML) potwierdziła również istnienie innego znanego powodu nieczytelności skryptów dowodowych. Jest nim dążenie do osiągnięcia konkretnego celu, jakim jest udowodnienie twierdzenia, przy jednoczesnym traktowaniu pobocznych wartości takich jak czytelność tworzonych skryptów dowodowych jako kwestie drugorzędne. Oczywiście efekt ten nasila się w przypadku długich rozumowań, które pojawiają się nieuchronnie w wyniku wzbogacania bazy MML w coraz trudniejsze zagadnienia. Niejednokrotnie obecne tam wywody, choć poprawne dla systemu weryfikującego są jednak mało eleganckie, często wręcz chaotyczne, a człowiek może je zrozumieć jedynie przy bardzo dużym nakładzie pracy i wysiłku. Czytelność takich skryptów wpływa nie tylko na czas ich analizowania, który jest potrzebny do wyodrębnienia idei sformalizowanych w ten sposób dowodów matematycznych, ale również na łatwość ich konserwacji (rewidowania). Autorzy takich skryptów mają przekonanie, iż problem znajdowania i usuwania niepotrzebnych lub możliwych do skrócenia fragmentów nie jest z ich punktu widzenia istotny, gdyż zakładają, że wszystkie elementy tego procesu można łatwo zautomatyzować. Przekonani są, że automatyczne odnajdywanie i usuwanie niepotrzebnych fragmentów skryptów dowodowych zapisanych w języku Mizar za pomocą narzędzi dystrybuowanych z tym systemem [12, 13] oraz wizualizacja częściowo poprawionych skryptów w zlinkowanej postaci HTML [19, 20] jest rozwiązaniem gwarantującym dostateczny poziom czytelności uzyskanych wywodów. Badając jednak dowody zgromadzone w bazie MML, bez trudu odnajdujemy jednopoziomowe rozumowania zbudowane z co najmniej 30 kroków, dla których rekonstrukcja idei nawet przy zastosowaniu dotychczas dostępnych narzędzi wymaga ogromnego nakładu pracy.

2 Obszar badawczy i cele rozprawy

Rozwiązaniem opisanego powyżej problemu mogą być metody stosowane w praktyce matematycznej w celu czytelnego formułowania długich rozumowań nieformalnych. Metody te są głęboko zakorzenione w tradycji formułowania dowodów matematycznych, jak również mają swoje uzasadnienie w wielu badaniach psychologicznych prowadzonych nad procesami poznawczymi człowieka [1, 2, 5]. Opierają się one na dwóch niezależnych od siebie środkach. Pierwszym z nich jest odnajdywanie i wyodrębnianie w postaci lematów lub kapsułkowaniu w postaci zagnieżdżonych podrozumowań mniej istotnych lub powtarzających się fragmentów rozumowania. Drugim zaś środkiem jest reorganizacja niezależnych od siebie kroków rozumowań w dowodzie, która ma na celu poprawę wybranych własności linearyzacji dowodu. Dodatkowo środki te są przenoszone na grunt dowodów formalnych zgromadzonych w bazie MML przez wielu autorów skryptów, dla których priorytetem w formalizacji jest nie tylko powiększanie bazy sformalizowanych twierdzeń, ale również jest jej jakość. Skrypty dowodowe tych autorów odznaczają się wówczas podwyższonym poziomem czytelności, a jeśli nawet zawierają niepotrzebne lub możliwe do skrócenia fragmenty, to ich usunięcie nie prowadzi do obniżenia czytelności tak zmodyfikowanych wywodów.

Naturalnie czytelność skryptów dowodowych jest pojęciem subiektywnym, różnie

rozumianym przez poszczególnych autorów rozumowań, a przedstawione metody są stosowane w sposób intuicyjny na podstawie wzorców występujących w dowodach nieformalnych, które zostały ukształtowane w wyniku wieloletniej praktyki ich tworzenia. W oparciu o te nieformalne wzorce, a także wieloletnie doświadczenia użytkowników systemu Mizar [15, 20] oraz badania nad czytelnością programów komputerowych [11] zaproponowano metody zwiększania czytelności sformalizowanych dowodów. Niestety nie wiadomo, na ile metody te są efektywne czasowo, a co za tym idzie, na ile są stosowalne w praktyce.

Wybór systemu Mizar został podyktowany tym, iż dystrybuowana z tym systemem biblioteka jest najbardziej rozbudowanym na świecie, intensywnie rozwijanym przez ponad dwadzieścia lat repozytorium sformalizowanej wiedzy matematycznej. Za wyborem bazy MML przemawia również fakt, że rozwiązania stosowane w języku Mizar wykorzystywane są do opracowywania podobnych rozwiązań służących polepszaniu czytelności skryptów w innych uznanych systemach [14], tj. *Declare* [18], *Mizar Mode for HOL* [9], *Isar language for Isabelle* [21], *Mizar-light for HOL-light* [23], *MMode for Coq* [8], *declarative proof language (DPL) for Coq* [4].

W związku z powyższym, jako **obszar badawczy** wybrano *wdrażanie metod wykorzystywanych w procesie poprawy czytelności nieformalnych rozumowań matematycznych w celu uczytelniania istniejących formalnych rozumowań zapisanych w systemie naturalnej dedukcji, ze szczególnym uwzględnieniem rozumowań zgromadzonych w bibliotece MML.*

Celem rozprawy w związku z tym jest *poprawa czytelności długich formalnych rozumowań poprzez upodobnienie ich do dowodów nieformalnych.* Cel ten był realizowany poprzez badanie efektywności algorytmów poprawiających czytelność rozumowań zapisanych w systemie naturalnej dedukcji G. Gentzena, S. Jaśkowskiego [7, 10].

W związku z tak nakreślonym celem głównym rozprawy zostały wyodrębnione następujące **cele szczegółowe** proponowanych badań:

1. *Wyodrębnienie zbioru informacji zawartych w poszczególnych krokach rozumowania, który umożliwia interpretowanie struktury rozumowania jako grafu skierowanego reprezentującego przepływ informacji między poszczególnymi krokami rozumowania, nazywanego dalej grafem dowodu.*

2. *Stworzenie abstrakcyjnego modelu grafu dowodu, który umożliwi uniezależnienie prowadzonych rozważań od konkretnego systemu weryfikującego, co uprości adaptację oczekiwanych rozwiązań do innych systemów.*

3. *Zdefiniowanie grupy wskaźników czytelności wyrażonych w terminach abstrakcyjnego grafu dowodu oraz zbadanie złożoności problemów grafowych optymalizujących wartości tych wskaźników.*

3 Problemy badawcze

Realizacja celów rozprawy nastąpiła przez odpowiedź na dwa sformułowane poniżej problemy badawcze:

1. Czy i w jakim stopniu możliwe jest posługiwanie się abstrakcyjnym modelem grafu dowodu do analizy metod uczytelniania istniejących rozumowań formalnych zgromadzonych w bazie MML?

2. Na ile proponowane metody uczytelniania są efektywne czasowo, a co za tym idzie, na ile są stosowalne w procesie automatycznej poprawy czytelności sformalizowanych rozumowań zapisanych w systemie naturalnej dedukcji?

Zdecydowano się na łączne potraktowanie tych dwóch problemów badawczych, ponieważ przeprowadzenie badań nad poprawą czytelności bez ich uniezależnienia od konkretnego systemu utrudniłoby, a wręcz uniemożliwiłoby adaptację oczekiwanych rozwiązań do innych systemów.

W rozprawie został przyjęty poniższy model abstrakcyjnego grafu.

Definicja 1. *Niech $\mathcal{V}$ będzie niepustym zbiorem, a $\mathcal{A}, \mathcal{M}$ będą dowolnymi podzbiorami zbioru uporządkowanych par elementów zbioru $\mathcal{V}$, $\mathcal{A}, \mathcal{M} \subseteq \mathcal{V} \times \mathcal{V}$. Uporządkowaną trójkę $\mathfrak{P} = \langle \mathcal{V}, \mathcal{A}, \mathcal{M} \rangle$ nazywamy* abstrakcyjnym grafem dowodu *wtedy i tylko wtedy, gdy:*

(*i*) *digraf $\mathfrak{M}_{\mathfrak{P}} = \langle \mathcal{V}, \mathcal{M} \rangle$ jest lasem dendroidów,*

(*ii*) *każdy łuk (u, v) w digrafie $\langle \mathcal{V}, \mathcal{A} \rangle$ spełnia warunek: każdy następnik u w lesie $\mathfrak{M}_{\mathfrak{P}}$ jest osiągalny z v i jednocześnie różny od v,*

(*iii*) *digraf $\langle \mathcal{V}, \mathcal{A} \cup \mathcal{M} \rangle$ jest acykliczny.*

Odnosząc się do pytania sformułowanego w pierwszym problemie badawczym, należy rozważyć dwa jego aspekty. Po pierwsze, należy zweryfikować, czy modyfikacja sposobu zlinearyzowania rozumowania poprzez modyfikację sortowań topologicznych poszczególnych rozumowań pierwotnych w grafie dowodu nie prowadzi do powstania błędów w skrypcie dowodowym. Po drugie, należy sprawdzić, czy rodziny abstrakcyjnych grafów dowodu, które zostały wykorzystane w rozprawie przy badaniu klasy złożoności problemów optymalizacji są konstruktywne, a więc czy rozumowania, których struktury są opisywane przez wykorzystywane grafy dowodów, należą lub potencjalnie mogą należeć do bazy MML.

Nawiązując do pierwszego aspektu problemu badawczego, należy stwierdzić, że zbiór informacji wyodrębnionych z kroków rozumowania sformułowanego w języku Mizar dostarcza koniecznej, jak również dostatecznej liczby danych do badania zależności występującej między krokami dowodu. Wynik ten został zbadany teoretycznie poprzez szczegółową analizę składni systemu Mizar jak również został potwierdzony empirycznie. Badania przeprowadzone nad wstępną poprawą budowy ponad 30 tys. rozumowań zgromadzonych w bazie MML, podczas których wykorzystywane narzędzia dokonywały modyfikacji sposobu linearyzacji rozumowań w oparciu jedynie o informacje wyodrębnione z kroków rozumowania, nie wygenerowały błędów w trakcie modyfikacji skryptów dowodowych. Stąd możemy wnioskować, że modyfikacja sposobu linearyzacji poszczególnych rozumowań pierwotnych przy uwzględnieniu jedynie informacje zgromadzonej w grafie dowodu, nie prowadzi do powstania błędów w ten sposób modyfikowanych skryptach dowodowych.

Odnosząc się do drugiego aspektu problemu badawczego, stwierdzamy, że wskazana w rozprawie podrodzina konstruktywnych abstrakcyjnych grafów dowodu nie pokrywa całej rodziny grafów dowodów rozumowań zgromadzonych w bazie MML, aczkolwiek zawiera wszystkie grafy spośród tych, które były wykorzystywane przy badaniu złożoności rozważanych problemów optymalizacyjnych. Naturalnie, teoretyczne zbadanie pełnej składni systemu Mizar doprowadziłoby do wyznaczenia zbioru własności abstrakcyjnych grafów dowodu dostatecznych do ich konstruktywności, aczkolwiek badanie to zostało uznane za niecelowe ze względu na realizację głównych celów rozprawy. Za dostateczne rozwiązanie tego problemu przyjęliśmy zostało bowiem przyjęte wskazanie transformacji, której zastosowanie powoduje, iż wynikowe dowody dają grafy będące konstruowanymi grafami dowodu.

Ustosunkowując się do powyższych wniosków w odniesieniu do pierwszego problemu badawczego stwierdzamy, że graf dowodu jest dobrym narzędziem umożliwiającym dostatecznie wierne odzwierciedlenie struktury rozumowań zapisanych w ję-

zyku Mizar. Skonstruowany model abstrakcyjnego grafu dowodu umożliwia nie tylko wierne odzwierciedlenie struktur rozumowań zapisanych w języku Mizar, ale również każdego rozumowania zapisanego w systemie naturalnej dedukcji. Własności przyjęte w definicji abstrakcyjnego grafu dowodu charakteryzują bowiem podstawowe zasady tworzenia poprawnie zbudowanych dowodów w systemie naturalnej dedukcji, tj.:

1) przesłanki, które są wykorzystywane w uzasadnieniu kroku, muszą być stwierdzone we wcześniejszej części rozumowania;

2) krok rozumowania, w którym stwierdzana formuła wykorzystuje zmienną wolną, musi być poprzedzony w rozumowaniu krokiem, w którym zmienna ta została ustalona w rozumowaniu;

3) uzasadnienie kroku nie może odwoływać się do stwierdzeń sformułowanych wewnątrz zagnieżdżonego rozumowania, jeśli krok ten nie należy do obszaru tego zagnieżdżonego rozumowania.

Możemy stąd wnioskować, że uzyskane wyniki rozprawy mogą być w prosty sposób zaadoptowane do każdego systemu opierającego się na systemie naturalnej dedukcji.

Odnosząc się do pytania sformułowanego w drugim problemie badawczym należy stwierdzić, że została zbadana złożoność problemu optymalizacji wszystkich pięciu najczęściej wskazywanych przez użytkowników systemu Mizar wskaźników czytelności, charakteryzujących następujące własności dowodu:

$\mathcal{K}.1$ Liczba kroków, z których każdy w swoim uzasadnieniu odwołuje się m.in. do przesłanki sformułowanej w bezpośrednio poprzedzającym kroku w dowodzie powinna być maksymalna.

$\mathcal{K}.2$ Liczba odwołań do przesłanek w obrębie poszczególnych liniowych fragmentów dowodu powinna być maksymalna.

$\mathcal{K}.3$ Liczba etykiet, które należy wprowadzić w dowodzie, w celu umożliwienia odwoływania się do daleko położonych przesłanek, które nie mogą być przekazane do uzasadnienia za pomocą konstrukcji `then` powinna być minimalna.

$\mathcal{K}.4$ Suma odległości po wszystkich odwołaniach między krokami, które odwołują się do przesłanek, a krokami, w których te przesłanki zostały uzasadnione powinna być minimalna.

$\mathcal{K}.5$ Liczba fragmentów rozumowania zapisanych spoiście w dowodzie, w których przepływ informacji jest dostatecznie gęsty, powinna być maksymalna.

Badania te przeprowadzono dwuetapowo. W pierwszym etapie zostały formalnie zdefiniowane wskaźniki opisujące linearyzacje poszczególnych rozumowań pierwotnych abstrakcyjnego grafu dowodu, jak również sformułowano pięć grafowych problemów decyzyjnych związanych z optymalizacją tych wskaźników:

$\mathcal{K}.1'$: Instancja: DAG $D = \langle \mathcal{V}, \mathcal{A} \rangle$, zbiór $A \subseteq \mathcal{A}$, liczba naturalna $0 \leqslant M_1 \leqslant |\mathcal{V}|$.

Pytanie: Czy istnieje acykliczna $\mathcal{H}_A^{(*)}$–partycja π digrafu D spełniająca zależność: $|\pi| \leqslant M_1$?

Przy czym partycja π digrafu D ma własność $\mathcal{H}_A^{(*)}$ wtedy i tylko wtedy, gdy dla każdego $p \in \pi$ podgraf digrafu $\langle \mathcal{V}, A \rangle$ indukowany przez zbiór wierzchołków P posiada drogę Hamiltona.

$\mathcal{K}.2'$: Instancja: DAG $D = \langle \mathcal{V}, \mathcal{A} \rangle$, zbiór $A \subseteq \mathcal{A}$, liczba naturalna $0 \leqslant M_2 \leqslant |A|$.
Pytanie: Czy istniej acykliczna $\mathcal{H}_A^{(*)}$–partycja π digrafu D spełniająca zależność: $\sum\limits_{\neq P_1, P_2 \in \pi} |P_1 \underset{A}{\curvearrowright} P_2| \leqslant M_2$?

Przy czym $P_1 \underset{A}{\curvearrowright} P_2$ oznacza zbiór $\{(v,u) \in A : v \in P_1 \wedge v \in P_2\}$.

$\mathcal{K}.3'$: Instancja: DAG $D = \langle \mathcal{V}, \mathcal{A} \rangle$, zbiory $A_2 \subseteq A_1 \subseteq \mathcal{A}$, liczba naturalna $0 \leqslant M_3 \leqslant |\mathcal{V}|$.
Pytanie: Czy istnieje sortowanie topologiczne $\tau \in TS(D)$ spełniające zależność:

$$|\{v \in \mathcal{V} : \underset{u \in \mathcal{V}}{\exists} vu \in A_1 \ \wedge \ (\, d_\tau(v,u) > 1 \ \vee \ vu \notin A_2 \vee \underset{w \in \mathcal{V}}{\exists} vw \in \mathcal{A} \setminus A_1 \,)\}| \leqslant M_3,$$

gdzie $d_\tau(v,u) = d(u) - d(v)$?

$\mathcal{K}.4'$: Instancja: DAG $D = \langle \mathcal{V}, \mathcal{A} \rangle$, funkcja $w : \mathcal{A} \to \mathbb{N} \cup \{0\}$, liczba naturalna $0 \leqslant M_4 \leqslant \binom{|\mathcal{V}|+1}{3} \cdot \max\limits_{vu \in \mathcal{A}} w(vu)$.
Pytanie: Czy istnieje sortowanie topologiczne $\tau \in TS(D)$ spełniające zależność:

$$\sum_{vu \in \mathcal{A}} w(vu) \cdot d_\tau(v,u) \leqslant M_4?$$

$\mathcal{K}.5'$: Instancja: DAG $D = \langle \mathcal{V}, \mathcal{A} \rangle$, zbiór $A \subseteq \mathcal{A}$, liczba rzeczywista $0 < p \leqslant 1$, liczba naturalna $0 \leqslant M_5 \leqslant 2^{|\mathcal{V}|}$.
Pytanie: Czy istnieje sortowanie topologiczne $\tau \in TS(D)$ spełniające zależność:

$$|\{V \subseteq \mathcal{V} : |V| \geqslant 2 \wedge \rho_A(V) \geqslant p \wedge V \text{ jest } \tau\text{–spoisty}\}| \geqslant M_5,$$

gdzie $\rho_A(V) = \dfrac{|\{\{v,u\} : v,u \in V \wedge v \neq u \wedge (v \underset{A}{\rightarrow}^* u \vee u \underset{A}{\rightarrow}^* v)\}|}{\binom{|V|}{2}}$, a zbiór V jest τ–spoisty wtedy i tylko wtedy, gdy istnieje liczba naturalna i, dla której $i \leqslant \tau(v) \leqslant i + |V| - 1$, dla każdego $v \in V$?

W drugim etapie została zbadana złożoność wszystkich sformułowanych uprzednio problemów. W tym celu w pierwszej kolejności udowodniono, że problem *Minimum Feedback Arc Set* (**FAS**, zob. GT8 [6]) zachowuje NP-zupełność jeśli nawet założymy dodatkowo o digrafie D występującym w instancji tego problemu, że każdy wierzchołek w D ma co najwyżej jednoelementowy zbiór następników lub poprzedników.

Minimum Feedback Arc Set (**FAS**)
Instancja: Digraf D, liczba naturalna $0 \leqslant K \leqslant |\mathcal{A}(D)|$.
Pytanie: Czy istnieje sprzężony podzbiór $\mathcal{A}(D)$–łuków o mocy co najwyżej K?

Następnie została skonstruowana redukcja wskazanego podproblemu **FAS** do pomocniczego problemu **APH** będącego podproblemem $\mathcal{K}.1$.

Acykliczna Partycja Hamiltona (**APH**)
Instancja: DAG $D = \langle \mathcal{V}, \mathcal{A} \rangle$ bez $\mathcal{A}$–skrótów, liczba naturalna $0 \leqslant K \leqslant |\mathcal{A}|$.
Pytanie: Czy istnieje acykliczna $\mathcal{H}_{\mathcal{A}}^{(*)}$–partycji digrafu D o mocy co najwyżej K?

Przy czym $\mathcal{A}$–skrótem nazywany $\mathcal{A}$–łuk, dla którego istnieje skierowana $\mathcal{A}$–ścieżka długości co najmniej 2 łącząca końce tego łuku.

Wykorzystanie pomocniczego problemu **APH** zostało podyktowane tym, iż problem ten redukuje się do problemu $\mathcal{K}.2$, co jest następstwem poniższego twierdzenia.

Twierdzenie 2. *Niech $D = \langle \mathcal{V}, \mathcal{A} \rangle$ będzie acyklicznym digrafem bez A–skrótów, M liczbą naturalną $1 \leqslant M \leqslant |\mathcal{V}|$, gdzie $A \subseteq \mathcal{A}$. Wówczas istnieje acykliczna $\mathcal{H}_A^{(*)}$–partycja π digrafu D o liczebności co najwyżej M wtedy i tylko wtedy, gdy istnieje acykliczna $\mathcal{H}_A^{(*)}$–partycja π digrafu D spełniająca zależność: $\sum_{\neq P_1, P_2 \in \pi} |P_1 \underset{A}{\curvearrowright} P_2| \leqslant |A| - |\mathcal{A}| + M$.*

NP-zupełny okazał się również problem $\mathcal{K}.4$, który w przypadku funkcji wagi przypisującej każdemu łukowi digrafu wartość 1 jest znanym NP-zupełnym problemem *Directed Optimal Linear Arrangement* (zob. GT43 [6]).

Ostatnim NP-zupełnym problemem spośród wskazanych pięciu okazał się problem $\mathcal{K}.5$. Wynik ten został uzyskany poprzez wskazanie redukcji problemu *Vertex Cover* (**VC**, zob. GT1 [6]) do $\mathcal{K}.5$, jeśli liczba rzeczywista p występująca w instancji problemu $\mathcal{K}.5$ ma wartość co najmniej $\frac{1}{3}$.

Vertex Cover (**VC**, GT1 [6]))
INSTANCJA: Graf nieskierowany $G = \langle \mathcal{V}, \mathcal{E} \rangle$, liczba naturalna $0 \leqslant K \leqslant |\mathcal{V}|$.
PYTANIE: Czy istnieje pokrycie wierzchołkowe $V \subseteq \mathcal{V}$ o mocy co najwyżej K?

Ograniczenie to nie pomniejsza jednak wartości uzyskanego wyniku, gdyż wielkości przyjmowane za ten parametr w procesie poprawy czytelności są na ogół bliskie jedności. Gęstość przepływu informacji o wartości 1 charakteryzuje bowiem te spośród podzbiorów kroków rozumowania, które można uporządkować tak, aby stanowiły liniowe fragmenty w zlinearyzowanym rozumowaniu.

Przeprowadzone badania wykazały również, że jedynym spośród sformułowanych problemów decyzyjnych, na który możemy udzielić odpowiedzi, stosując algorytm o złożoności wielomianowej jest problem $\mathcal{K}.3$. Na szczególną uwagę zasługuje fakt, że wynik ten jest następstwem jedynie syntaktycznych ograniczeń na stosowalność konstrukcji `then` w systemie Mizar, która to umożliwia przekazanie do uzasadnienia przesłanki, jeśli jest ona sformułowana w bezpośrednio poprzedzającym kroku, bez podania explicite etykiety wskazującej na tę przesłankę.

Pominięcie tego ograniczenia skutkuje bowiem modyfikacją pytania sformułowanego w problemie $\mathcal{K}.3'$ do postaci występującej w problemie $\mathcal{K}.3''$, który jest problemem NP-zupełnym, co zostało udowodnione rozprawie.

$\mathcal{K}.3''$: INSTANCJA: DAG $D = \langle \mathcal{V}, \mathcal{A} \rangle$, zbiory $A_2 \subseteq A_1 \subseteq \mathcal{A}$, liczba naturalna $0 \leqslant M_3 \leqslant |\mathcal{V}|$.
PYTANIE: Czy istnieje sortowanie topologiczne $\tau \in TS(D)$ spełniające zależność:

$$|\{v \in \mathcal{V} : \underset{u \in \mathcal{V}}{\exists} vu \in A_1 \wedge (d_\tau(v,u) > 1 \vee vu \notin A_2)\}| \leqslant M_3?$$

Przedstawione wyniki badań utwierdzają zatem w przekonaniu, że poprawa czytelności rozumowań zapisanych w języku Mizar wiąże się w większości przypadków z rozwiązywaniem NP-zupełnych problemów grafowych. Stąd wnioskujemy, że aplikacje, które mogłyby realizować poprawę czytelności w zadowalającym czasie, są w stanie co najwyżej aproksymować optymalne wartości wskaźników. Dodatkowo, przeprowadzone wstępne badania empiryczne pokazały również, że zbiory sortowań

topologicznych, w których zaproponowane w rozprawie wskaźniki mają wartości optymalne w ogólnym przypadku nie zawierają elementów wspólnych. Stąd wnioskujemy, że aplikacje uczytelniające skrypty dowodowe mogą działać jedynie przy ustalonej hierarchii wartości optymalizowanych wskaźników.

Wstępne badania przeprowadzone nad metodami poprawy czytelności wykorzystującymi wyodrębnianie fragmentów rozumowania w postaci lematów (nazywane dalej *paczkami*) umożliwiły uzupełnienie odpowiedzi na pytanie sformułowane w drugim problemie badawczym. Wykazały one bowiem, że możliwe jest wyodrębnianie z rozumowania dowolnych paczek przy zachowaniu poprawności modyfikowanych skryptów dowodowych, wykorzystując do tego celu algorytm o złożoności wielomianowej [16]. Przeprowadzone badania umożliwiły zatem uzyskanie istotnego wyniku, który w przyszłości będzie przydatny w określeniu kierunku dalszych analiz dotyczących narzędzi, których zadaniem będzie automatyczne odnajdywanie i wyodrębnianie fragmentów rozumowania. Wstępne wyniki badań sugerowały bowiem, iż stwierdzenie opisujące rozumowanie zawarte w paczce powinno mieć postać implikacji, której poprzednikiem jest koniunkcja przesłanek, które jednocześnie są zlokalizowane poza obszarem paczki oraz są wykorzystane w krokach paczki, natomiast następnikiem – koniunkcja stwierdzeń uzasadnionych w paczce, które są wykorzystywane w dalszej części rozumowania poza paczką. Następstwem takiego uproszczenia w sposobie opisu rozumowania zawartego w paczce są cykle skierowane powstające w modyfikowanym grafie dowodu, jeśli wyodrębniane paczki nie są domknięte na prowadzenie dróg skierowanych. Natomiast ograniczenie możliwości wyodrębniania jedynie do przypadku, w którym paczki są domknięte na prowadzenie dróg skierowanych zawężało przestrzeń poszukiwań partycji rozumowania na paczki do partycji acyklicznych. Ograniczenie to sprowadzało więc problem podziału rozumowania do rozwiązywania znanego NP-zupełnego problemu *Acyclic Partition* (zob. ND15 [6]). Stworzenie metody wyodrębniania nawet niedomkniętych paczek stanowi zatem istotny punkt w badaniach nad poprawą czytelności przy wykorzystaniu wyodrębniania paczek. Niestety wyodrębnienie takich paczek jest bardziej skomplikowane, aniżeli w przypadku paczek domkniętych oraz wiąże się z powielaniem części kroków rozumowania. Stąd możemy wnioskować, że algorytm poszukujący najbardziej optymalnej partycji rozumowania na paczki, choć nie będzie zawężał swoich poszukiwań jedynie do partycji acyklicznych, to jednak będzie musiał uwzględniać jako jeden z optymalizowanych wskaźników liczbę „wychodzących" dróg skierowanych w paczkach, które to odpowiadają za powstawanie cykli w grafie partycji.

4 Wnioski

Sformułowana odpowiedz na pytanie postawione w drugim problemie badawczym potwierdza więc pierwszą część sformułowanej na wstępie **hipotezy rozprawy**, którą jest następujące stwierdzenie: *nie jest możliwe osiągnięcie wartości optymalnej większości proponowanych wyznaczników poprawy czytelności stosując algorytmy o złożoności wielomianowej. Nie jest również możliwe równoczesne osiągnięcie wartości optymalnych wszystkich proponowanych wyznaczników. Istnieją jednak algorytmy, umożliwiające optymalizację wielokryterialną wyznaczników poprawy czytelności, które działając przy ustalonej hierarchii ważności zaproponowanych metod, dostosowanej do potrzeb określonej grupy czytelników, mogą skutecznie poprawić czytelność analizowanych rozumowań formalnych.* Druga część hipotezy została potwierdzona za pomocą empirycznych badań przeprowadzonych w ramach „wstępnej" poprawy budowy rozumowań. Zaproponowany algorytm poszukujący linearyzacji rozumowania przy

ustalonej hierarchii ważności wyznaczników czytelności umożliwił bowiem zestandaryzowanie rozumowań zgromadzonych w bazie MML. Wykorzystywanie algorytmu zachłannego nie prowadzi jednak w ogólnym przypadku do odnajdywania rozwiązań aproksymujących z zadaną dokładnością rozwiązanie optymalne. Analiza zmodyfikowanych skryptów dowodowych wykazała jednak, że zastosowanie algorytmu dokonującego lokalnej optymalizacji okazało się skutecznym narzędziem umożliwiającym odnajdywanie i spoisty zapis w zlinearyzowanym rozumowaniu „spójnych" podrozumowań.

Do efektywnej poprawy czytelności i jasności zgromadzonych rozważań niezbędne jest jednak prowadzenie dalszych badań nad algorytmami dokonującymi optymalizacji wielokryterialnej, jak również nad opracowaniem standardów czytelności dowodów.

Bibliografia

[1] O. Behaghel, *Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern*, Indogermanische Forschungen, t. 25, s. 110–142, 1909.

[2] D. E. Broadbend, *Decision and Stress*, London Academic Press, 1971.

[3] N. G. de Bruijn, *A Survey of the Project Automath*, w: J. P. Seldin i in. (edt.), To H.B. Curry: Essays in Combinatory Logic, Lambda Calculus, and Formalism, s. 589–606, Academic Press, 1980.

[4] P. Corbineau, *A Declarative Language for the Coq Proof Assistant*, w: Proc. of the 2007 International Conference on Types for Proofs and Programs, LNCS, t. 4941, s. 69–84, 2007.

[5] N. Cowan, *Attention and Memory: An Integrated Framework*, Oxford University Press, 1998.

[6] M. R. Garey, D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, A Series of Books in the Mathematical Science, 1979.

[7] G. Gentzen, *Untersuchungen über das logische Schließen*, Mathematische Zeitschrift, t. 35, nr 1, s. 176–210, 1935.

[8] M. Giero, F. Wiedijk, *MMode, A Mizar Mode for the proof assistant Coq*, ICIS, Radboud Universiteit Nijmegen, 2004.

[9] J. Harrison, *A Mizar Mode for HOL*, w: Proc. of the 9th International Conference on Theorem Proving in Higher Order Logics, LNCS, t. 1125, s. 203–220, Springer-Verlag, 1996.

[10] S. Jaśkowski, *On the Rules of Supposition in Formal Logic*, Studia Logica, 1934.

[11] A. J. Ko, B. A. Myers, M. J. Coblenz, Htet Htet Aung, *An Exploratory Study of How Developers Seek, Relate, and Collect Relevant Information during Software Maintenance Tasks*, IEEE Transactions On Software Engineering, t. 32, nr 12, s. 971–988, 2006.

[12] R. Milewski, *New Auxiliary Software for MML Database Management*, Mechanized Mathematics and Its Applications, t. 5, nr 2, s 1–10, 2006.

[13] R. Milewski, *Transformations of MML Database's Elements*, w: Proc. of Mathematical Knowledge Management 2006, LNCS, t. 3863, s. 376–388, Springer-Verlag, 2006.

[14] A. Naumowicz, A. Korniłowicz, *A Brief Overview of Mizar*, w: S. Berghofer i in. (red.), Theorem Proving in Higher Order Logics, LNCS, t. 5674, s. 67–72, Springer-Verlag, 2009.

[15] K. Pąk, *The Algorithms for Improving and Reorganizing Natural Deduction Proofs*, Studies in Logic, Grammar and Rhetoric, t. 22, nr 35, s 95–112, 2010.

[16] K. Pąk, *Methods of Lemma Extraction in Natural Deduction Proofs*, Journal of Automated Reasoning, t. 50, nr 2, s. 217–228, 2013.

[17] E. Snapper, *The Three Crises in Mathematics: Logicism, Intuitionism and Formalism*, Mathematics Magazine, t. 52, nr 4, s. 207–216, 1979.

[18] D. Syme, *Three Tactic Theorem Proving*, w: Theorem Proving in Higher Order Logics, LNCS, t. 1690, s. 203–220, Springer-Verlag, 1999.

[19] J. Urban, *MizarMode - An Integrated Proof Assistance Tool for the Mizar Way of Formalizing Mathematics*, Journal of Applied Logic, t. 4, nr 4, s. 414–427, 2006.

[20] J. Urban, *XML-izing Mizar: Making Semantic Processing and Presentation of MML Easy*, w: Proc. of Mathematical Knowledge Management 2005, LNCS, t. 3863, s. 346–360, Springer-Verlag, 2006.

[21] M. Wenzel, F. Wiedijk, *A Comparison of Mizar and Isar*, Journal of Automated Reasoning, t. 29, nr 3–4, s. 389–411, 2002.

[22] A. N. Whitehead, B. Russell, *Principia Mathematica*, Cambridge Mathematical Library, Cambridge University Press, 1910.

[23] F. Wiedijk, *Mizar Light for HOL Light*, w: J. B. Richard i in. (red.), Proc. of the 14th International Conference on Theorem Proving in Higher Order Logics, LNCS, t. 2152, s. 378–393, Springer-Verlag, 2001.

[24] F. Wiedijk, *The De Bruijn Factor*.