

Autoreferat pracy doktorskiej

Paweł Gora

14.06.2023

Niniejszy dokument zawiera autoreferat mojej pracy doktorskiej zatytułowanej “Metaheuristics in Optimization of Complex Processes” (“Metaheurystyki w optymalizacji procesów złożonych”).

1 Motywacja badań

Badania systemów złożonych w ujednoczonych ramach zostały uznane za nową, interdyscyplinarną dziedzinę naukową [5, 37]. Zajmuje się ona systemami składającymi się z wielu oddziałujących na siebie obiektów, nie w pełni odseparowanych od otoczenia. Te obiekty i ich interakcje powodują dynamiczne, skomplikowane, nieliniowe i często chaotyczne zachowania. Takie systemy nazywane są *systemami złożonymi*, podczas gdy ich ewolucje w czasie nazywane są *procesami złożonymi*.

Pojęcie systemów i procesów złożonych może obejmować struktury abstrakcyjne, np. automaty komórkowe, a także struktury fizyczne, które mogą występować w wielu obszarach, od poziomu mikroskopowego opisującego stan każdego pojedynczego obiektu, do makroskopowego opisującego ogólny stan systemu. Przykładami fizycznych procesów złożonych są ruch drogowy w miastach i żywe organizmy.

Kompleksowe zrozumienie funkcjonowania procesu złożonego i zidentyfikowanie metod optymalizacji jego ewolucji może być bardzo korzystne. Zwykle jednak bardzo trudno jest kontrolować zachowanie takich systemów ze względu na liczbę składników i charakter ich interakcji, które wprowadzają nieliniowość, czułą zależność od warunków początkowych i obliczeniową nieredukowalność. Otwartość systemu pozwala dodatkowo na interakcje z innymi obiektami, co może prowadzić do niedeterminizmu. Obecność takich cech zwykle wskazuje, że zachowanie danego systemu może być nieprzewidywalne oraz trudne do zrozumienia i kontroli. Jednocześnie, sterowanie procesami złożonymi jest bardzo ważne z praktycznego punktu widzenia, ponieważ optymalizacja procesów złożonych jest krytycznym aspektem wielu

branż. Usprawniając operacje, firmy i instytucje publiczne mogą zwiększyć produktywność, obniżyć koszty i poprawić zadowolenie użytkowników. Dlatego ważne jest badanie tego obszaru i próba opracowania uniwersalnych metod analizy i optymalizacji procesów złożonych.

Pomimo wielu podobieństw, istnieją zasadnicze różnice między różnymi systemami złożonymi, a ogólny aparat matematyczny nie jest wystarczająco dojrzały, aby badać wszystkie możliwe systemy na raz. Dlatego moja praca doktorska koncentruje się na 2 procesach złożonych:

1. Głównym badanym procesem jest ruch pojazdów w miastach, jest on opisany w Sekcji 4.2. Jest to proces składający się z wielu samochodów/kierowców, których celem jest bezpieczne i jak najszybsze dotarcie do celu. Samochody muszą współdzielić infrastrukturę, więc wchodzi w interakcje, co może prowadzić do korków lub wypadków drogowych. Istnieje wiele sposobów optymalizacji ruchu drogowego, ale niniejsza rozprawa skupia się wyłącznie na sterowaniu sygnalizacją świetlną.
2. Proces ewolucji nowotworu pod wpływem radioterapii, opisany w Sekcji 4.3. W tym przypadku żywe komórki oddziałują zgodnie z określonymi zasadami, a niektóre z nich mogą stać się komórkami nowotworowymi. Radioterapia może wpłynąć na ten proces i ostatecznie zmniejszyć liczbę komórek nowotworowych.

Jak omówiono w rozprawie, optymalizacja obu procesów jest bardzo ważna dla społeczeństwa i może pomóc w rozwiązaniu istotnych problemów gospodarczych i cywilizacyjnych, takich jak zatępienie komunikacyjne, wypadki samochodowe, a także cierpienie i śmierć z powodu nowotworu.

Aby zastosować niektóre techniki optymalizacji, konieczne jest posiadanie modelu matematycznego procesu złożonego. Jak wyjaśniono w Rozdziale 4, istnieje wiele podejść do modelowania procesów złożonych i zazwyczaj nie ma uniwersalnej metody, która zawsze jest najlepsza. Wszystkie modele mają zalety i wady, więc wybór najlepszych modeli powinien zależeć od zamierzonego zastosowania i celu modelowania.

Rozdział 4 zawiera przegląd najważniejszych aspektów modelowania procesów złożonych i przedstawia przykłady modeli, skupiając się na procesach złożonych badanych w niniejszej rozprawie: ruchu drogowym w miastach (Sekcja 4.2) i ewolucji nowotworu w trakcie radioterapii (Sekcja 4.3).

Gdy model procesu złożonego jest już dostępny, można go wykorzystać do oszacowania wartości funkcji celu, która powinna zostać zoptymalizowana. Problemy optymalizacyjne wynikające z procesów złożonych są zazwyczaj niewypukłe i często posiadają wiele minimów lokalnych. Przeprowadzono

już wiele badań nad optymalizacją procesów złożonych, np. [7, 34]. Jednak większość z nich koncentruje się na konkretnych procesach złożonych i nadal pozostaje kwestią otwartą, czy istnieją uniwersalne, skuteczne techniki optymalizacji wszystkich lub szerokiego zakresu procesów złożonych.

Wśród popularnych technik, które można uznać za uniwersalne, są metaheurystyki, takie jak algorytmy genetyczne [30] lub optymalizacja rojem cząstek [23]. Są to algorytmy optymalizacyjne wykorzystywane do znajdowania przybliżonych rozwiązań złożonych problemów optymalizacyjnych, jakościowo bliskich rozwiązaniom optymalnym. W przeciwieństwie do tradycyjnych algorytmów optymalizacyjnych, które opierają się na wzorach matematycznych w celu znalezienia optymalnego rozwiązania, metaheurystyki wykorzystują heurystykę do kierowania poszukiwaniami rozwiązań w zazwyczaj bardzo dużej przestrzeni możliwości. To sprawia, że dobrze nadają się do rozwiązywania problemów, w których optymalne rozwiązanie jest trudne lub niemożliwe do określenia przy użyciu klasycznych modeli matematycznych. Metaheurystyki są również często stosowane, gdy problem optymalizacyjny jest zbyt duży lub zbyt złożony, aby można go było rozwiązać dokładnymi metodami w rozsądnym czasie.

Aby formalnie określić metaheurystykę, konieczne jest zdefiniowanie funkcji celu (funkcji dopasowania) oceniającej jakość rozwiązań, a także kryterium stopu określającego kiedy algorytm powinien przestać szukać lepszego rozwiązania (np. maksymalna liczba iteracji, minimalny próg poprawy) oraz reguł przejścia od bieżącego rozwiązania do nowego rozwiązania, np. przy użyciu losowości lub technik adaptacyjnych w celu zrównoważenia eksploracji i eksploatacji.

Wspomniane eksploracja i eksploatacja to dwa podstawowe pojęcia w kontekście metaheurystyk. Dobry algorytm metaheurystyczny powinien zachować równowagę między eksploracją i eksploatacją, aby zapewnić skuteczne przeszukiwanie przestrzeni rozwiązań i znajdowanie rozwiązań dobrej jakości.

Metaheurystyki nie gwarantują znalezienia optymalnego rozwiązania, ale często są w stanie w rozsądnym czasie znaleźć rozwiązania zbliżone do optymalnych. W kontekście procesów złożonych metaheurystyka może być wykorzystywana do znajdowania dobrych wartości parametrów kontrolujących ewolucję procesu. Badania w tym obszarze są już od wielu lat przeprowadzane, a przegląd najpopularniejszych metaheurystyk, które są również rozważane w tej rozprawie, znajduje się w Rozdziale 6.

W niektórych prowadzonych przeze mnie wstępnych badaniach nad rozwiązywaniem problemu ustawiania sygnalizacji świetlnej, zdefiniowanego w Sekcji 8.1.1 jako “Traffic Signal Setting problem”, metaheurystyki (zwłaszcza algorytmy genetyczne) dały obiecujące wyniki [17], ale czas obliczeń był dość długi, co wykluczało sterowanie sygnalizacją świetlną w czasie rzeczy-

wistym, a nawet poważnie utrudniało proces przeprowadzania eksperymentów i walidacji różnych metaheurystyk z różnymi ustawieniami. To skłoniło mnie do zbadania w jaki sposób metaheurystyki mogą być zastosowane do rozwiązywania problemu ustawiania sygnalizacji świetlnej oraz optymalizacji innych procesów złożonych w rozsądnym czasie obliczeniowym. Moim głównym celem nadal był problem ustawiania sygnalizacji świetlnej przy użyciu metaheurystyk, ale naturalnym celem badawczym było opracowanie technik, które mogłyby być na tyle uniwersalne, aby można je było zastosować do optymalizacji wielu innych procesów złożonych.

Rozproszenie obliczeń na wysokowydajnych klastrach obliczeniowych oraz uruchamianie eksperymentów w infrastrukturze chmurowej nie dało wystarczającego przyspieszenia, więc konieczne było poszukiwanie innych technik. Ostatecznie pokonałem problemy obliczeniowe, konstruując metodę optymalizacji opartą na połączeniu metaheurystyk z modelami surogatywnymi mającymi na celu zastąpienie początkowego, trudnego obliczeniowo modelu, modelami uczenia maszynowego, które mogą być ewaluowane dużo szybciej.

Doprowadziło to do pojawienia się kilku nowych pytań badawczych, takich jak:

1. Które modele uczenia maszynowego są najlepszymi modelami surogatywnymi dla mikroskopowych symulacji ruchu drogowego?
2. Które metaheurystyki dadzą najlepsze wyniki i są najbardziej efektywne w rozwiązywaniu problemu ustawiania sygnalizacji świetlnej w połączeniu z modelami surogatywnymi bazującymi na uczeniu maszynowym?
3. Czy ta technika jest na tyle uniwersalna, że można ją zastosować do optymalizacji innych procesów złożonych? Jeśli tak, to jakie są najlepsze modele surogatywne i metaheurystyki dla danego procesu złożonego? Czy zależy to od charakteru procesu złożonego?

W celu znalezienia odpowiedzi na te pytania przeprowadziłem wiele serii eksperymentów, które wykazały wydajność różnych modeli surogatywnych i metaheurystyk w rozwiązywaniu problemu ustawiania sygnalizacji świetlnej, a następnie zweryfikowałem wprowadzoną metodologię, stosując ją do optymalizacji innego procesu złożonego – ewolucji nowotworu podczas leczenia radioterapią.

2 Cele badawcze

Głównym celem badań było głębsze zrozumienie natury procesów złożonych i zrobienie kroku w kierunku opracowania uniwersalnych, skalowalnych metod analizy i optymalizacji procesów złożonych przy użyciu metaheurystyk. W rozprawie skupiłem się na dwóch procesach złożonych, ale pewne bardziej ogólne rozważania dotyczące procesów złożonych są również zawarte w tej pracy, m.in. w Rozdziałach 3 i 7. Pierwszym i podstawowym badanym procesem złożonym był miejski ruch drogowy sterowany sygnalizacją świetlną. Celem było zoptymalizowanie niektórych ważnych charakterystyk ruchu (przede wszystkim całkowitego czasu oczekiwania pojazdów na czerwonym świetle w danym obszarze i w danym przedziale czasu) za pomocą metaheurystyk i zbadanie, które z nich dają najlepsze wyniki pod względem jakości rozwiązań i skalowalności, w tym czasu obliczeń. Drugim badanym procesem złożonym była ewolucja nowotworu podczas leczenia radioterapią, a celem było zminimalizowanie liczby komórek nowotworowych po leczeniu poprzez kontrolowanie wielkości dawek i terminów podawania radioterapii.

Początkowe eksperymenty wykazały, że połączenie metaheurystyk z modelami symulacyjnymi może pomóc znajdować dość dobre ustawienia sygnalizacji świetlnej, ale jest to kosztowne obliczeniowo ze względu na potrzebę oceny dużej liczby potencjalnych rozwiązań za pomocą modeli symulacyjnych. Złożoność czasowa uniemożliwiła sterowanie sygnalizacją świetlną w czasie rzeczywistym. W rzeczywistości bardzo trudno było przeprowadzić eksperymenty i zweryfikować różne metaheurystyki z różnymi ustawieniami, nawet jeśli obliczenia były wykonywane na klastrze wielu procesorów. Duża złożoność obliczeniowa realistycznych symulacji, ich walidacja i wykorzystanie jako ewaluatorów funkcji dopasowania dla metaheurystyk stały się dla mnie poważnym problemem, który musiałem rozwiązać, aby przetestować wiele algorytmów optymalizacji z różnymi ustawieniami.

Dlatego moim kolejnym celem badawczym było sprawdzenie, czy metaheurystyki mogą być zastosowane do optymalizacji procesów złożonych w rozsądnym czasie obliczeniowym. Wysiłki mające na celu przyspieszenie obliczeń doprowadziły do opracowania metodologii optymalizacji bazującej na połączeniu metaheurystyk i modelowania surogatywnego z wykorzystaniem uczenia maszynowego.

Metoda ta dała obiecujące wyniki w rozwiązywaniu problemu ustawiania sygnalizacji świetlnej, więc kolejnym celem badawczym było sprawdzenie, które modele uczenia maszynowego są najlepszymi modelami surogatywnymi dla mikroskopowych symulacji ruchu drogowego, a także które metaheurystyki mogą dać najlepsze wyniki i są najbardziej wydajne w rozwiązywaniu problemu ustawiania sygnalizacji świetlnej w połączeniu z modelami suroga-

tywnymi bazującymi na uczeniu maszynowym.

Opracowana metoda wydawała się być na tyle uniwersalna, by mogła zostać zastosowana do optymalizacji innych procesów złożonych, więc kolejnym celem badawczym było przetestowanie jej w innym przypadku użycia: optymalizacji ewolucji nowotworu podczas radioterapii. W tym przypadku celem było zminimalizowanie liczby komórek nowotworowych po kilku dniach leczenia poprzez kontrolowanie terminów i wielkości dawek radioterapii.

3 Metodologia

Metody badania optymalizacji procesów złożonych opracowane w ramach prezentowanych badań były testowane w środowiskach symulacyjnych, tak zwanych *cyfrowych bliźniakach* rozważanych procesów złożonych.

W przypadku sterowania sygnalizacją świetlną podstawowym narzędziem był opracowany przeze mnie program do symulacji ruchu drogowego – Traffic Simulation Framework (TSF). Podstawowy model ruchu został również zaproponowany przeze mnie jako rozszerzenie modelu Nagela-Schreckenberga (NaSch) [31] do przypadku realistycznych sieci drogowych. Zarówno model, jak i jego implementacja, zostały opisane w Rozdziale 5 niniejszej rozprawy.

Dla symulacji wzrostu nowotworów opracowano inne narzędzie, bazujące na modelu opisanym w [2]. Prace nad jego implementacją zostały zaproponowane, zaprojektowane i nadzorowane przeze mnie, ale przeprowadzone przez grupę studentów w ramach projektu programistycznego związanego z ich pracą licencjacką [4].

W dalszej części badań narzędzia symulacyjne dla badanych procesów złożonych zostały zastosowane do ewaluacji funkcji celu: wartości niektórych parametrów symulacji zostały przekazane jako dane wejściowe, a symulatory obliczały wartości funkcji celu odpowiadające danym metrykom, np. całkowity czas oczekiwania na czerwonych światłach, całkowity czas podróży, całkowita liczba komórek nowotworowych. Następnie tak skonstruowane narzędzia ewaluacji zostały użyte w badanych algorytmach optymalizacyjnych.

W pracach przedstawionych w niniejszej rozprawie przeanalizowano kilka metaheurystyk, np. algorytmy genetyczne, optymalizację roju cząstek, symulowane wyżarzanie, przeszukiwanie tabu, strategię ewolucji adaptacji macierzy kowariancji (CMA-ES), algorytmy memetyczne i optymalizację bayesowską. Wszystkie te algorytmy zostały opisane w Rozdziale 6.

Ze względu na znaczne problemy obliczeniowe i długi czas potrzebny do oceny jakości ustawień sterowania procesami złożonymi, konieczne było znalezienie sposobu na przyspieszenie obliczeń. W tym celu wykorzystano modele surogatywne bazujące na technikach uczenia maszynowego, głównie sieciach

neuronowych i drzewach decyzyjnych ze wzmocnieniem gradientowym (np. LightGBM). Główną ideą takich modeli surogatywnych było zastąpienie czasochłonnych symulacji procesów złożonych w celu uzyskania przybliżonych wyników symulacji. Modele surogatywne bazujące na uczeniu maszynowym zostały wytrenowane i przetestowane na zbiorach danych wygenerowanych przy użyciu symulacji komputerowych. Po wytrenowaniu możliwe było zintegrowanie ich z metaheurystykami i innymi algorytmami optymalizacyjnymi w celu przyspieszenia oceny jakości potencjalnych rozwiązań znajdujących podczas eksploracji dużych przestrzeni możliwych opcji (ustawień sterowania procesów złożonych).

W późniejszych eksperymentach jakość najlepszych znalezionych rozwiązań została oceniona przez oryginalny model symulacyjny, aby sprawdzić, czy modele surogatywne są w stanie przybliżać prawidłowe wyniki symulacji z wystarczającą dokładnością. Przeprowadzone eksperymenty miały na celu przetestowanie średnich bezwzględnych błędów procentowych (MAPE) oraz maksymalnych bezwzględnych błędów procentowych (MAXAPE) modeli surogatywnych na zestawach testowych i w pobliżu rozwiązań znalezionych za pomocą metaheurystyk.

Modele surogatywne zawsze oferują kompromis między dokładnością a wydajnością obliczeniową, dlatego ważne jest, aby wybrać takie modele, które zapewnią odpowiednią dokładność i czas obliczeń. We wstępnych eksperymentach testowałem również modele surogatywne bazujące na innych technikach uczenia maszynowego, takich jak maszyny wektorów wspierających (SVM) [10], oraz na mezoskopowym modelu ruchu (pośrednim między modelami mikroskopowymi i makroskopowymi), ale podejścia te nie dały zadowalających wyników. Późniejsze próby zaprojektowania lepszych modeli surogatywnych, które podjąłem wspólnie z koordynowaną przeze mnie grupą badawczą TensorCell, zaowocowały opracowaniem nowej architektury rzadkich grafowych sieci neuronowych, a także zbadaniem wpływu technik łączenia kilku modeli, czy stosowania różnych funkcji aktywacji i funkcji strat. Okazało się, że obiecującym podejściem może być np. Margin Ranking Loss zaprojektowany w celu zachowania porządku aproksymowanych ustawień zgodnego z porządkiem wyznaczonym przez oryginalny, przybliżony model. Przeprowadzono również wiele innych eksperymentów w celu zbadania różnych modeli surogatywnych opartych na uczeniu maszynowym. Były one związane m.in. z aktywnym uczeniem się (modele surogatywne były dotrenowywane przy użyciu punktów bliskich dobrym rozwiązaniem znalezionym za pomocą metaheurystyki) i technikami wyjaśnialności modeli uczenia maszynowego. Znaczny wysiłek włożono również w badanie dokładności modeli surogatywnych w pobliżu punktów uznanych przez nie za bliskie optimum lokalnych znalezionych za pomocą algorytmów optymalizacyjnych.

Zaproponowane podejście do optymalizacji procesów złożonych, łączące symulacje komputerowe, metaheurystyki i modele surogatywne, zostało przeze mnie sformalizowane w Rozdziale 7 i zbadane w eksperymentach przedstawionych w Rozdziale 8. Łącznie podczas tych badań przeprowadzono wiele serii eksperymentów, a rozprawa zawiera opis najważniejszych kamieni milowych - jest to 10 eksperymentów związanych ze sterowaniem sygnalizacją świetlną, 2 eksperymenty związane z optymalizacją leczenia nowotworu za pomocą radioterapii, a także 1 bardziej ogólny eksperyment mający na celu zbadanie, jak trudno jest aproksymować stany automatów komórkowych za pomocą modeli uczenia maszynowego. Były to następujące eksperymenty:

1. Eksperymenty z algorytmami genetycznymi i TSF (Sekcja 8.3.1.1).
2. Eksperymenty z algorytmami genetycznymi i modelem mezoskopowym (Sekcja 8.3.1.2).
3. Trenowanie w pełni połączonych sieci neuronowych typu feed-forward jako modeli surogatywnych dla TSF (Sekcja 8.3.1.3).
4. Badanie różnych modeli sieci neuronowych i strategii ich trenowania (Sekcja 8.3.1.4).
5. Zastosowanie w pełni połączonych modeli sieci neuronowych typu feed-forward do optymalizacji ustawień sygnalizacji świetlnej (Sekcja 8.3.1.5).
6. Badanie wydajności sieci neuronowych i modeli gradient boosting jako modeli surogatywnych (Sekcja 8.3.1.6).
7. Optymalizacja ustawień sygnalizacji świetlnej przy użyciu metody spadku gradientowego (Sekcja 8.3.1.7).
8. Testowanie różnych metaheurystyk i modeli surogatywnych (Sekcja 8.3.1.8).
9. Eksperymenty z grafowymi sieciami neuronowymi jako modelami surogatywnymi (Sekcja 8.3.1.9).
10. Końcowe eksperymenty z różnymi metaheurystykami i modelami surogatywnymi (Sekcja 8.3.1.10).
11. Eksperymenty z symulatorem wzrostu nowotworu i algorytmami genetycznymi (Sekcja 8.3.2.1).
12. Eksperymenty z modelami surogatywnymi i metaheurystykami (Sekcja 8.3.2.2).

13. Przybliżanie automatów komórkowych (Sekcja 8.3.3).

Eksperymenty te miały na celu zbadanie jakości różnych metaheurystyk i modeli surogatywnych w rozważanych zadaniach. Kolejne eksperymenty doprowadzały każdorazowo do wniosków, które zostały zastosowane w następnych eksperymentach. Umożliwiły one również zidentyfikowanie ograniczeń wprowadzonej metodologii, które omówiono w Rozdziale 9, wraz z kilkoma pomysłami na ich przewyżczenie w celu zbliżenia się do praktycznych zastosowań opracowanej metody.

4 Dyskusja wyników

Przeprowadzone eksperymenty wykazały, że zastąpienie kosztownych obliczeniowo symulacji modelami surogatywnymi może być bardzo przydatne i po integracji z algorytmami optymalizacyjnymi może prowadzić nie tylko do znacznego przyspieszenia eksperymentów, testowania większej liczby algorytmów, hiperparametrów i potencjalnych rozwiązań, ale także do znalezienia lepszych rozwiązań oryginalnego problemu optymalizacyjnego. W eksperymentach przetestowano wiele różnych modeli uczenia maszynowego, metaheurystyk i ich ustawień w wielu seriach eksperymentów, a kolejne eksperymenty były projektowane na podstawie wyników wcześniejszych. Co ważne, metodologia okazała się skuteczna w przypadku obu badanych procesów złożonych - ruchu drogowego w miastach i ewolucji nowotworu podczas radioterapii.

W przypadku problemu ustawiania sygnalizacji świetlnej, po przeprowadzeniu wielu systematycznych eksperymentów, dla różnych obszarów, z różnymi modelami surogatywnymi, różnymi algorytmami optymalizacji i z różnymi wartościami hiperparametrów, nie można jednoznacznie stwierdzić, które modele surogatywne, algorytmy optymalizacyjne i ich parametry są najlepsze. Może to zależeć od wielu czynników. Techniki *populacyjne* (takie jak algorytmy genetyczne, algorytmy memetyczne i CMA-ES) wydają się jednak dawać lepsze wyniki niż metody *indywidualne* (takie jak symulowane wyżarzanie), choć np. metoda spadku gradientowego również wydaje się być dość efektywna.

Porównując modele surogatywne okazuje się, że architektury rzadkich grafowych sieci neuronowych zaproponowane podczas badań i modele LightGBM wydają się działać z podobną jakością i lepiej niż np. modele w pełni połączonych sieci typu feed-forward. Jedną z zalet grafowych sieci neuronowych jest to, że można je stosować w połączeniu z metodą spadku gradientowego, która jest stosunkowo szybka i, zgodnie z przeprowadzonymi eksperymentami, wymaga najmniejszej liczby wywołań modelu surogatywnego

w celu uzyskania zbieżności. Jednak spadek gradientowy zwykle zbiega do lokalnego optimum, więc rozwiązania znalezione przy użyciu tej metody są zazwyczaj gorsze niż rozwiązania znalezione przy użyciu metaheurystyk *populacyjnych*, takich jak CMA-ES lub algorytmy memetyczne. Algorytmy te wymagają większej liczby wywołań modelu, aby uzyskać zbieżność, ale ostatecznie są w stanie znaleźć nieco lepsze (według mikroskopowego modelu w programie TSF) ustawienia sygnalizacji świetlnej. Z drugiej strony, różnice w jakości nie są znaczne, a może to być również spowodowane własnościami optymalizowanych funkcji celu i lokalizacją ich minimów w rozważanej przestrzeni możliwych rozwiązań.

Większość testowanych algorytmów optymalizacji osiągnęła dobre wyniki i w rozsądnym czasie znalazła ustawienia sygnalizacji świetlnej znacznie lepsze niż najlepsze ustawienia w losowo generowanych treningowych zbiorach danych, więc wydaje się, że ogólna metodologia jest skuteczna. Zalecenia dotyczące stosowania konkretnych modeli surogatywnych i metaheurystyk mogą zależeć od konkretnego przypadku użycia i priorytetów, ponieważ zazwyczaj będzie to kompromis między jakością pożądanых rozwiązań, a czasem potrzebnym na ich znalezienie.

W przypadku optymalizacji leczenia nowotworu za pomocą radioterapii eksperymenty wykazały, że dzięki przyspieszeniu symulacji, przeprowadzeniu większej liczby eksperymentów z algorytmem genetycznym, i z większą liczbą konfiguracji hiperparametrów, możliwe było znalezienie lepszych protokołów radioterapii. Eksperymenty wykazały również, że symulator wzrostu nowotworu pod wpływem radioterapii można z powodzeniem zastąpić modelami surogatywnymi bazującymi na uczeniu maszynowym w zadaniu znajdowania dobrych protokołów radioterapii przy użyciu algorytmów genetycznych. Wyniki są nie tylko tej samej jakości, ale również czas przeprowadzania eksperymentów jest o kilka rzędów wielkości krótszy. Oczywiście początkowo wygenerowanie zbioru danych i wytrenowanie modeli wymagało znacznego czasu, ale gdy model surogatywny jest gotowy, daje to możliwość zbadania jeszcze większego zestawu potencjalnych protokołów radioterapii. Dlatego też metodologia zaproponowana w tej rozprawie wydaje się być bardzo przydatna.

Innym interesującym wnioskiem było to, że techniki łączenia wielu modeli (ensemble learning), a także łączenie sieci neuronowych z mechanizmem uwagi [41] i uwzględnianie Margin Ranking Loss [32] jako funkcji straty może również poprawić dokładność modeli surogatywnych. Techniki te mogą potencjalnie skutkować projektowaniem lepszych modeli surogatywnych w przyszłości, co zostało również podkreślone w Rozdziale 9.

Ponieważ wprowadzona metodologia okazała się przydatna w znajdowaniu dobrych ustawień sterowania innym procesem złożonym (oprócz ruchu

drogowego w miastach), naturalne było zbadanie, czy ta sama metodologia może być stosowana w przypadku innych procesów złożonych i czy istnieją jakieś ograniczenia tej metodologii.

W przypadku obu procesów badanych w tej rozprawie modele matematyczne bazowały na automatach komórkowych, a wytrenowane modele surogatywne oparte na uczeniu maszynowym były w stanie aproksymować wyniki tych modeli z bardzo dobrą dokładnością i przy stosunkowo niewielkich zbiorach treningowych i czasach uczenia. Osiągnięcie dobrej dokładności aproksymacji jest gwarantowane przez *Twierdzenie o uniwersalnej aproksymacji* [12], ale to, czy i w jakich okolicznościach można to robić efektywnie, może zależeć od wielu czynników. Dlatego naturalne było rozważenie w jakich przypadkach automaty komórkowe można łatwo aproksymować za pomocą modeli uczenia maszynowego. Temat ten jest szczególnie interesujący biorąc pod uwagę fakt, że niektóre automaty komórkowe są Turing-zupełne, co oznacza, że są w stanie wykonać dowolne obliczenia, które może wykonać maszyna Turinga. Jednym z najbardziej znanych przykładów automatów komórkowych Turing-zupełnych jest słynny automat “Gra w życie” (“Game of Life”) zaproponowany przez J. Conwaya [15, 36]. Najprostszym automatem komórkowym, o którym wiadomo, że jest Turing-zupełny jest automat znany jako “Reguła 110” (“Rule 110”) [42, 11].

Postanowiłem zbadać jak trudno jest wytrenować sieci neuronowe do przewidywania stanów “Reguły 110”, “Gry w Życie”, a także deterministycznego wariantu oryginalnego modelu NaSch [31]. Eksperymenty zostały przeprowadzone przez grupę studentów Uniwersytetu Warszawskiego pod moim nadzorem, a wyniki zostały podsumowane w ich pracy licencjackiej [35].

Okazuje się, że dla automatów Turing-zupełnych (“Reguła 110” i “Gra w Życie”) przewidywanie przyszłych stanów jest trudniejsze niż dla modelu NaSch. Ponadto, biorąc pod uwagę, że takie automaty komórkowe można efektywnie zaimplementować, a często również przyspieszyć za pomocą procesorów graficznych, nie ma dużej przewagi pod względem czasu obliczeń dla małej liczby kroków, a w przypadku przewidywania stanów po większej liczbie kroków dokładność jest gorsza.

Jest to zgodne z wynikami opublikowanymi niedawno w pracy [38], gdzie stwierdzono, że rozmiar sieci neuronowych wymaganych do nauczenia się funkcji wejścia/wyjścia reprezentowanej przez wiele kroków “Gry w życie” jest często znacznie większy niż minimalna sieć wymagana do zaimplementowania tej funkcji.

Analiza powiązanych prac naukowych pokazała, że ten interesujący temat nie jest wciąż do końca zbadany. Przypuszczam, że dalsze badania mogą doprowadzić do eksperymentalnych, a także teoretycznych, np. związanych z teorią automatów, wyników pokazujących trudność aproksymacji stanów

różnych automatów komórkowych, w szczególności Turing-zupełnych, za pomocą modeli uczenia maszynowego, co może mieć również wpływ na potencjalną stosowalność metodologii wprowadzonej w niniejszej rozprawie. Z drugiej strony, wydaje się również, że dla niektórych automatów komórkowych, np. deterministycznego wariantu modelu Nagela-Schreckenberga lub jego rozszerzenia w narzędziu TSF, łatwiej jest przewidzieć przy pomocy uczenia maszynowego pewne własności ewolucji automatu.

5 Podsumowanie efektów prac i oryginalnego wkładu

Badania przeprowadzone i przedstawione w niniejszej rozprawie zaowocowały następującym wkładem naukowym:

1. Dalsze rozwinięcie na przestrzeni wielu lat programu Traffic Simulation Framework [18], w szczególności do postaci narzędzia oceniającego jakość różnych ustawień sygnalizacji świetlnej przy użyciu mikroskopowego modelu ruchu. Można go uznać za “cyfrowego bliźniaka” rzeczywistego ruchu miejskiego. Może on służyć do oceny jakości różnych ustawień sterowania ruchem.
2. Opracowanie nowego oprogramowania do symulacji ewolucji nowotworów i oceny jakości protokołów radioterapii. Zaproponowałem i zaprojektowałem rozwój tego nowego narzędzia w oparciu o model opisany w pracy [2], a zostało ono zaimplementowane przez grupę studentów, których nadzorowałem [4]. Dzięki zastosowaniu języka C++ i zrównolegleniu obliczeń przy pomocy kart graficznych, narzędzie to znacznie przewyższa poprzednie implementacje pod względem szybkości. Program ten można uznać za “cyfrowego bliźniaka” żywej tkanki z komórkami nowotworowymi poddawanych radioterapii.
3. Nowa definicja systemów złożonych, procesów złożonych i ich modeli, zaproponowana na podstawie analizy niektórych innych definicji istniejących w literaturze naukowej.
4. Ogólna metodologia optymalizacji procesów złożonych oparta na połączeniu symulacji komputerowych, metaheurystyk i modeli surogatywnych bazujących na uczeniu maszynowym.
5. Zastosowanie modeli uczenia maszynowego jako modeli surogatywnych, które przybliżają wyniki czasochłonnych symulacji komputerowych. We-

dług mojej wiedzy, było to pierwsze zastosowanie modeli uczenia maszynowego jako modeli surogatowych do symulacji ruchu drogowego kontrolowanego poprzez sygnalizację świetlną, a także do symulacji ewolucji nowotworu podczas radioterapii. Metodologia ta przezwyciężyła wyzwania obliczeniowe napotkane podczas badań i umożliwiła przetestowanie szerokiego spektrum metaheurystyk w różnych ustawieniach. Pomysł został zaproponowany przeze mnie, a implementacja i eksperymenty zostały przeprowadzone we współpracy z naukowcami z grupy badawczej TensorCell, którą założyłem i prowadzę od samego jej początku. Przebadalem wiele różnych modeli surogatowych w celu określenia ich wydajności w zależności od wielu czynników.

6. Opracowanie nowej architektury rzadkich grafowych sieci neuronowych, w których topologia połączeń bazuje na topologii rozważanej sieci drogowej. Oczekuję, że architektura ta znajdzie zastosowanie również w innych zadaniach związanych z ruchem drogowym i problemami grafowymi. Architektura została początkowo zaproponowana przez jednego z członków grupy badawczej TensorCell, Łukasza Skowronka, ale ja koordynowałem całość badań i eksperymentów.
7. Kompleksowa ocena kilku metaheurystyk zastosowanych do problemów optymalizacyjnych związanych z dwoma procesami złożonymi: ruchem drogowym w mieście i ewolucją nowotworu.
8. Wygenerowanie dużej liczby zbiorów danych przy użyciu symulatorów ruchu drogowego i ewolucji nowotworów. Zbiory te zostały opublikowane, aby ułatwić społeczności naukowej dalsze badania nad tymi tematami.

Jednym z efektów przedstawionych prac było również założenie przeze mnie w 2017 roku niezależnej grupy badawczej TensorCell [40] w celu prowadzenia badań nad optymalizacją procesów złożonych. Chociaż nie jest to wynik naukowy, można go uznać za wartościowy efekt prac, ponieważ kilkoro członków zespołu zainteresowało się domeną optymalizacji procesów złożonych przy użyciu sztucznej inteligencji i istotnie rozwinęło swoje umiejętności. Co więcej, zespół przygotował kilka wartościowych publikacji naukowych, opublikowanych na dobrych konferencjach informatycznych (np. NeurIPS) i transportowych (np. MT-ITS) i można się spodziewać dalszego wartościowego wkładu naukowego tej grupy w przyszłości.

6 Wkład rozprawy w kontekście prac pokrewnych

Jak przedstawiono w Rozdziale 2 rozprawy, w literaturze naukowej istnieje już wiele podejść do optymalizacji procesów złożonych, w tym sterowania sygnalizacją świetlną i optymalizacji leczenia nowotworu, a także do modelowania surogatywnego. Metody stosowane do optymalizacji procesów złożonych, które można uznać za raczej uniwersalne, obejmują m.in. metaheurystyki [13, 20] i uczenie ze wzmocnieniem [39, 3, 14], a także interaktywne obliczenia granularne [21]. W przypadku modelowania surogatywnego najpopularniejszymi technikami są “response surfaces” [6], “kriging” [25], podejścia bayesowskie [8] i modele uczenia maszynowego [10, 1]. Rozdział 4 pokazuje z kolei, że dziedzina modelowania procesów złożonych jest również bardzo szeroka.

Rozdział 2 i niektóre inne sekcje rozprawy przedstawiają jednak również zidentyfikowane słabości istniejących technik optymalizacji procesów złożonych, które są stosowane w praktyce, w tym brak proaktywności (przewidywania przyszłych warunków i odpowiedniego dostosowywania ustawień z wyprzedzeniem), brak możliwości dokładnej oceny wprowadzanych zmian, wątpliwa wydajność w dużej skali, ograniczona adaptowalność do zmieniającego się środowiska i stosunkowo mała przestrzeń możliwych modyfikacji. Wszystkie techniki stosowane do optymalizacji procesów złożonych mają pewne mocne i słabe strony, wymagają więc znajdowania kompromisów między różnymi aspektami. Na przykład, zwiększenie jakości oceny wprowadzonych zmian lub zwiększenie rozmiaru przestrzeni możliwych działań zwykle prowadzi do wyższych kosztów obliczeniowych. Dlatego nadal istnieje potrzeba dalszego badania i rozwijania różnych technik. Na podstawie przeglądu aktualnego stanu wiedzy doszedłem do wniosku, że kwestia optymalizacji procesów złożonych może być nadal uważana za otwarty problem, dla którego nowe rozwiązania są bardzo pożądane przez naukowców i inżynierów.

Zgodnie z moją wiedzą, podejście bazujące na wykorzystaniu modeli uczenia maszynowego do aproksymacji wyników symulacji ruchu drogowego opisane w rozprawie jest pierwszym takim podejściem w zagadnieniu sterowania sygnalizacją świetlną, zwłaszcza w połączeniu z metaheurystykami. Pomysł ten został po raz pierwszy zaprezentowany w 2016 roku na warsztatach “NIPS 2016 Workshop on Nonconvex Optimization for Machine Learning: Theory and Practice” i nawet recenzenci podkreślali, że warto zbadać zastosowanie sieci neuronowych jako modeli surogatywnych w tym kontekście. Możliwe, że było to również jedno z pierwszych takich podejść w całej dziedzinie inżynierii ruchu drogowego. Architektury rzadkich grafowych sieci neu-

ronowych wprowadzone jako modele surogatywne symulacji ruchu są również nowością w tym kontekście.

Jeśli chodzi o wykorzystanie metaheurystyk, publikacje naukowe na temat ich zastosowań do sterowania sygnalizacją świetlną lub leczenia nowotworu są stosunkowo nowe, co wskazuje, że podejścia te w tych dziedzinach są również stosunkowo nowoczesne. Badania przedstawione w niniejszej rozprawie były jednymi z pierwszych takich podejść do systematycznego porównania wielu różnych metaheurystyk do sterowania sygnalizacją świetlną. Podobne badanie (ale tylko z 3 metaheurystykami, przy użyciu bardzo prostego i mniej realistycznego modelu ruchu) zostało przedstawione w [9].

Jak przedstawiono w Rozdziale 4, istnieje już wiele podejść do modelowania procesów złożonych, w tym ruchu drogowego i rozwoju nowotworów, ale Traffic Simulation Framework, który był jednym z głównych narzędzi wykorzystanych w tym badaniu, był również jednym z pierwszych rozwiązań stosujących mikroskopowe modele ruchu w skali dużych miast. Rozszerzenie przeze mnie modelu NaSch [31] pozwoliło na przeprowadzenie wydajnych symulacji ruchu na realistycznej sieci drogowej Warszawy, co było znaczącym osiągnięciem w czasach, gdy narzędzie to powstawało (do dziś większość symulatorów ruchu drogowego działających w dużej skali stosuje również mniej dokładne modele mezoskopowe lub makroskopowe). Jest to również jeden z powodów, dla których program ten stał się popularny i znalazł zastosowanie w innych pracach badawczych (por. Rozdział 5).

Nowatorskim podejściem wydaje się również zastosowanie modeli surogatywnych do aproksymacji wyników symulacji rozwoju nowotworu w celu optymalizacji radioterapii przy pomocy metaheurystyk. W przypadku modelu rozwoju nowotworu wykorzystanego w niniejszej rozprawie (por. Sekcja 4.3), opracowanie nowego narzędzia symulacyjnego korzystającego z kart graficznych również dało znaczącą przewagę i pozwoliło na zbadanie większej liczby ustawień algorytmów optymalizacji, a także ocenę większej liczby potencjalnych rozwiązań. W efekcie udało się dzięki temu znaleźć lepsze protokoły radioterapii.

Co równie ważne, moje podejście polegające na badaniu dwóch procesów złożonych z pozornie zupełnie różnych dziedzin (inżynierii ruchu i medycyny) celem zbudowania uniwersalnych metod zdolnych do optymalizacji procesów złożonych z wielu różnych dziedzin jest również dość wyjątkowe. Naukowcy zajmujący się transportem lub medycyną rzadko postrzegają badane przez siebie procesy jako szczególne przypadki pewnych bardziej ogólnych zjawisk. Takie ogólne podejście jest bardziej powszechne i naturalne dla badaczy zajmujących się systemami złożonymi, ale nawet dla nich połączenie wiedzy i technik z kilku różnych dziedzin (takich jak automaty komórkowe, symulacje komputerowe, uczenie maszynowe i metaheurystyki) jest zwykle wyzwaniem.

W kontekście zidentyfikowanych ograniczeń istniejących technik optymalizacji procesów złożonych, metody przedstawione w niniejszej rozprawie mają na celu zmniejszenie kosztów obliczeniowych przy jednoczesnym zapewnieniu proaktywności i zdolności do dokładnej oceny wpływu wprowadzanych ustawień, biorąc pod uwagę stosunkowo dużą przestrzeń możliwych opcji. Metody te nadal nie są bezpośrednio adaptowalne do zmieniającego się środowiska, więc w tym celu należy raczej stosować inne techniki i jest to również jeden z powodów, dla których poza metaheurystykami i modelowaniem surogatywnym badałem również zastosowania uczenia ze wzmocnieniem, co omówiłem w Rozdziale 9 (istotny potencjał mogą mieć również techniki transfer learning [43]).

7 Wnioski

Cele określone na początku badań zawartych w niniejszej rozprawie (przedstawione w Sekcji 1.2) okazały się wyzwaniem, ale ostatecznie udało się je osiągnąć. Zrozumiałem głębiej naturę procesów złożonych, podsumowałem swoje badania w tym zakresie w Rozdziałach 4 i 7. Poszerzyłem również swoją wiedzę i zdobyłem dużo doświadczeń w innych obszarach, takich jak metaheurystyki, uczenie maszynowe, inteligentne systemy transportowe, modelowanie matematyczne, obliczenia naukowe, optymalizacja kombinatoryczna i obliczenia kwantowe. Ta nowa wiedza i umiejętności już zaprocentowały w badaniach podsumowanych w niniejszej rozprawie, a także w innych projektach badawczych, w które byłem zaangażowany.

Jednym z efektów było zaproponowanie uniwersalnej metody optymalizacji procesów złożonych (co było drugim celem prezentowanych badań, por. Rozdział 1.2), poprzez połączenie symulacji komputerowych, metaheurystyk i modeli surogatywnych bazujących na uczeniu maszynowym. Metoda ta okazała się wystarczająco dobra pod względem złożoności obliczeniowej i jakości znajdowanych rozwiązań, umożliwiła przeprowadzenie wielu serii eksperymentów. Możliwe było przetestowanie tej metody dla dwóch zastosowań pochodzących z pozornie zupełnie różnych dziedzin: sterowania sygnalizacją świetlną dla ruchu drogowego w miastach i optymalizacji radioterapii w leczeniu nowotworu.

Zbadano kilka metaheurystyk i porównano je z innymi algorytmami optymalizacyjnymi i okazało się, że ogólnie metaheurystyki *populacyjne* (algorytmy genetyczne, algorytmy memetyczne i CMA-ES) dają najlepsze wyniki, ale niektóre inne algorytmy (np. spadek gradientowy) są również dobre.

Zbadano także kilka modeli surogatywnych bazujących na sieciach neuronowych i LightGBM, a także zaproponowano nową architekturę rzadkich

grafowych sieci neuronowych, która okazała się dawać dobre przybliżenie wyników symulacji ruchu drogowego. Dla aproksymacji wyników symulacji ewolucji nowotworu pod wpływem radioterapii LightGBM również okazał się dawać zadowalające wyniki, ale jeszcze lepsze rezultaty uzyskano dla splątowanych (konwolucyjnych) sieci neuronowych z mechanizmem uwagi. Dalszą poprawę uzyskano dzięki łączeniu kilku modeli i zastosowaniu nowej funkcji straty (Margin Ranking Loss).

Pomimo początkowych trudności obliczeniowych, udało mi się przeprowadzić wiele serii eksperymentów. Większość wygenerowanych zbiorów danych oraz niektóre programy i kody źródłowe zostały udostępnione publicznie, aby ułatwić dalsze prace nad tym tematem.

W Rozdziale 2 przedstawiłem kilka innych dobrze znanych technik optymalizacji procesów złożonych i budowania modeli surogatowych oraz omówiłem ich ograniczenia, a także wkład metod wprowadzonych w niniejszej rozprawie w kontekście powiązanych prac. Z drugiej strony, w Rozdziale 9 omówiłem również zidentyfikowane ograniczenia proponowanej metody i pomysły na sposoby ich przezwyciężenia. Zaprezentowałem również możliwości dalszego rozszerzenia metody w celu zastosowania jej w praktyce.

Warto podkreślić, że aktualnie trwają dyskusje na temat zastosowania rozważanej metody w rzeczywistych systemach zarządzania ruchem oraz wykorzystania Traffic Simulation Framework w niedawno utworzonym Smart-City Lab w Chełmie. Ponieważ oprogramowanie TSF zostało opracowane na podstawie sieci drogowej i danych o ruchu z pomiarów w Warszawie, istnieje również potencjał do przeprowadzenia pilotażowych projektów i zaplanowania możliwych wdrożeń w tym mieście. Pojawił się również pomysł zastosowania opracowanych w ramach niniejszej rozprawy technik w nowej dziedzinie - inżynierii materiałowej.

Założyłem grupę badawczą TensorCell [40], której celem jest zastosowanie technik bazujących na sztucznej inteligencji do optymalizacji procesów złożonych i rozwiązywania problemów optymalizacji kombinatorycznej. Grupa ta przygotowała już kilka artykułów naukowych i zaprezentowała je na najważniejszych konferencjach dotyczących transportu i sztucznej inteligencji (np. MT-ITS, NIPS / NeurIPS).

Badania objęte niniejszą rozprawą zostały podsumowane i opublikowane łącznie w 25 publikacjach naukowych oraz nagrodzone kilkoma prestiżowymi nagrodami, w tym nagrodami "LIDER ITS" za najlepszą pracę badawczo-rozwojową w dziedzinie inteligentnych systemów transportowych w Polsce (w 2015 i 2017 roku [26, 27]), "Top 10 Polish Talents" MIT Technology Review [29] oraz "New Europe 100" [33].

8 Struktura pracy

Praca ma następującą strukturę:

- Rozdział 1 zawiera wprowadzenie, przedstawienie tematu, motywacji badań i celów badawczych. Znajduje się w nim również krótkie przedstawienie metodologii omawianej w pracy oraz spis efektów prac badawczych i struktura dalszej części pracy.
- Rozdział 2 zawiera kompleksowy przegląd istniejącej literatury i wiedzy w dziedzinach najbardziej istotnych dla metodologii i podejść badawczych przedstawionych w niniejszej rozprawie.
- Rozdział 3 koncentruje się na wyjaśnieniu koncepcji systemów złożonych i procesów złożonych oraz analizuje ich podstawowe cechy.
- Rozdział 4 zawiera kompleksowy przegląd metod modelowania procesów złożonych istotnych z punktu widzenia niniejszej rozprawy.
- Rozdział 5 zawiera prezentację programu Traffic Simulation Framework, który był jednym z narzędzi opracowanych przez mnie podczas prezentowanych badań i jest jednym z wyników prac.
- Rozdział 6 wyjaśnia czym są metaheurystyki oraz zawiera przegląd najpopularniejszych algorytmów metaheurystycznych, które zostały później wykorzystane w eksperymentach opisanych w rozprawie.
- Rozdział 7 wyjaśnia ogólną metodologię optymalizacji procesów złożonych przy użyciu metaheurystyk i modeli surogatowych, które zostały zaproponowane w niniejszej rozprawie.
- Rozdział 8 przedstawia główne eksperymenty przeprowadzone w ramach badań wraz z omówieniem ich wyników.
- Rozdział 9 zawiera omówienie zidentyfikowanych ograniczeń wprowadzonej metody, możliwych sposobów ich przewyciężenia, a także potencjalnych rozszerzeń i przyszłych kierunków badań, które mogą prowadzić do rzeczywistych zastosowań.
- Rozdział 10 podsumowuje rozprawę.

Po tych rozdziałach znajduje się Bibliografia oraz Dodatek z wykazem dostępnych zbiorów danych (Dodatek A) i programów (Dodatek B).

Literatura

- [1] C. Angione, E. Silverman, and E. Yaneske, “Using machine learning as a surrogate model for agent-based simulations”. *PLOS ONE*, vol. 17, no. 2, 2022. DOI: 10.1371/journal.pone.0263150.
- [2] S.D. Angus and M.J. Piotrowska, “A Matter of Timing: Identifying Significant Multi-Dose Radiotherapy Improvements by Numerical Simulation and Genetic Algorithm Search”. *PLoS ONE*, vol. 9, no. 12, e114098, 2014. DOI: doi.org/10.1371/journal.pone.0114098.
- [3] J. Ault and G. Sharon, “Reinforcement Learning Benchmarks for Traffic Signal Control”, *Proceedings of the Thirty-fifth Conference on Neural Information Processing Systems (NeurIPS 2021) Datasets and Benchmarks Track*, 2021.
- [4] R. Banaś, J. Bazińska, O. Łobożewicz, and A. Strzałka, “Optimization of cancer treatment”, Bachelor thesis in Computer Science at University of Warsaw, supervised by P. Gora, 2019.
- [5] Y. Bar-Yam, “Dynamics of Complex Systems”, Westview Press, 1997. DOI: 10.1201/9780429034961.
- [6] G.E.P. Box and K.B. Wilson, “On the Experimental Attainment of Optimum Conditions”, *Journal of the Royal Statistical Society, Series B*, vol. 13, no. 1, pp. 1–45, 1951. DOI: 10.1007/978-1-4612-4380-9_23.
- [7] L. Bussolari, P. Contucci, C. Giardina, C. Giberti, F. Unguendoli, and C. Vernia, “Optimization Strategies in Complex Systems”, arXiv: math/0309058, 2003.
- [8] I.C. Cardenas, “On the use of Bayesian networks as a meta-modeling approach to analyse uncertainties in slope stability analysis”, *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*, vol. 13, no. 1, pp. 53–65, 2019. DOI: 10.1080/17499518.2018.1498524.
- [9] S.W. Chen, C.B. Yang, and Y.H. Peng, “Algorithms for the Traffic Light Setting Problem on the Graph Model”, *Proceedings of the 12th Conference on Artificial Intelligence and Applications*, TAAI 2007.
- [10] A. Ciccazzo, G.D. Pillo and V. Latorre, “Support vector machines for surrogate modeling of electronic circuits”, *Neural Computing & Applications*, vol. 24, pp. 69–76, 2014. DOI: 10.1007/s00521-013-1509-5.

- [11] M. Cook, “Universality in Elementary Cellular Automata”, *Complex Systems*, vol. 15, no. 1, pp. 1–40, 2004. DOI: 10.25088/ComplexSystems.15.1.1.
- [12] G. Cybenko, “Approximation by superpositions of a sigmoidal function”, *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, 1989. DOI: 10.1007/BF02551274.
- [13] K.F. Doerner, M. Gendreau, P. Greistorfer, W. Gutjahr, R.F. Hartl, and M. Reimann, “Metaheuristics: Progress in Complex Systems Optimization”, Springer, 2007. DOI: 10.1007/978-0-387-71921-4.
- [14] S. El-Tantawy, B. Abdulhai, and H. Abdelgawad, “Multiagent Reinforcement Learning for Integrated Network of Adaptive Traffic Signal Controllers (MARLIN-ATSC): Methodology and Large-Scale Application on Downtown Toronto”, in *IEEE Conference on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1140-1150, 2013. DOI: 10.1109/TITS.2013.2255286.
- [15] M. Gardner, “Mathematical games: The fantastic combinations of John Conway’s new solitaire game ”life””, *Scientific American*, vol. 223, no. 4, pp. 120-123, 1970.
- [16] M. Gendreau and J.-Y. Potvin., “*Handbook of Metaheuristics*”, Springer, 2010. DOI: 10.1007/978-1-4419-1665-5.
- [17] P. Gora, “Adaptive planning of vehicular traffic”, M.Sc. thesis in Computer Science, University of Warsaw, 2010.
- [18] P. Gora, “Traffic Simulation Framework - a Cellular Automaton based tool for simulating and investigating real city traffic”, *Recent Advances in Intelligent Information Systems*, pp. 641-653, 2009.
- [19] P. Gora and K. Kurach, “Approximating Traffic Simulation using Neural Networks and its Application in Traffic Optimization”, ‘*NIPS 2016 Workshop on Nonconvex Optimization for Machine Learning: Theory and Practice*, 2016.
- [20] A. Jamal, H.M. Al-Ahmadi, F.M. Butt, M. Iqbal, M. Almoshaogeh, and S. Ali, “Metaheuristics for Traffic Control and Optimization: Current Challenges and Prospects”, *Search Algorithm - Essence of Optimization*, 2023. DOI: 10.5772/intechopen.99395.

- [21] A. Jankowski, A. Skowron, and M. Szczuka, “Interactive Granular Computing in Rightly Judging Systems”, *Rough Sets and Knowledge Technology*, vol. 5589, pp. 1–16, 2009. DOI: 10.1007/978-3-642-02962-2_1.
- [22] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, et. al., “LightGBM: A Highly Efficient Gradient Boosting Decision Tree”, *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [23] J. Kennedy and R. Eberhart, “Particle Swarm Optimization”, *Proceedings of IEEE International Conference on Neural Networks*, vol. IV, pp. 1942–1948, 1995. DOI: 10.1109/ICNN.1995.488968.
- [24] D.G. Krige, “A statistical approach to some mine valuations and allied problems at the Witwatersrand”, Master’s thesis at the University of Witwatersrand, 1951.
- [25] K.R. Krohn and J.L. Rhodes, “Algebraic theory of machines”, *Proceedings of the Symposium on Mathematical Theory of Automata*, 1962.
- [26] “LIDER ITS 2015”, results of the contest, <https://www.itspolska.pl/2015/05/25/wynik-vi-edycji-konkursu-lider-its-2015>. Last accessed: 27.05.2023.
- [27] “LIDER ITS 2017”, results of the contest, <https://www.itspolska.pl/2017/05/16/wyniki-viii-edycji-konkursu-lider-its-2017>. Last accessed: 27.05.2023.
- [28] S. Luke, “*Essentials of Metaheuristics*”, 2nd edition, 2013, available at <http://cs.gmu.edu/~sean/book/metaheuristics>. Last accessed: 10.06.2023.
- [29] MIT Technology Review, information about the announcement of the Top 10 Polish Talents of 2017 in the MIT Innovators Under 35 competition, <https://linktopoland.com/en/mit-technology-review-announces-top-10-polish-talents-2017>. Last accessed: 27.05.2023.
- [30] M. Mitchell, “An Introduction to Genetic Algorithms”, Cambridge, MA, MIT Press, 1996. DOI: 10.7551/mitpress/3927.001.0001.
- [31] K. Nagel and M. Schreckenberg, “A cellular automaton model for freeway traffic”, *Journal de Physique I*, vol. 2, no. 12, pp. 2221–2229, 1992. DOI: 10.1051/jp1:1992277.

- [32] M. Nayyeri, X. Zhou, S. Vahdati, H.S. Yazdi, and J. Lehmann, “Adaptive Margin Ranking Loss for Knowledge Graph Embeddings via a Correntropy Objective Function”, arXiv:1907.05336, 2019.
- [33] New Europe 100 list 2017, information about the recognitions, <https://businessinsider.com.pl/rozwoj-osobisty/kariera/new-europe-100-nagrody-dla-innowatorow-nominacje-2017/pmkhqfm>. Last Accessed: 27.05.2023.
- [34] T. Pham Dinh, “Optimization of Complex Systems: Theory, Models, Algorithms and Applications”, *Advances in Intelligent Systems and Computing*, vol. 991. Springer, 2020. DOI: 10.1007/978-3-030-21803-4.
- [35] R. Pragacz, P. Senchanka, N. Siwek N., and S. Zwara, “Approximation of cellular automata simulations using machine learning algorithms”, Bachelor thesis at the University of Warsaw, supervised by P. Gora, 2019.
- [36] P. Rendell, “Turing machine universality of the Game of Life”, *Genet Program Evolvable Mach*, vol. 18, pp. 115–117, 2017. DOI: 10.1007/s10710-016-9284-6.
- [37] A.F. Siegenfeld and Y. Bar-Yam, “An Introduction to Complex Systems Science and Its Applications”, *Complexity*, vol. 2020, Article ID 6105872, 2020. DOI: 10.1155/2020/6105872.
- [38] J.M. Springer and G.T. Kenyon, “It’s Hard for Neural Networks to Learn the Game of Life”, *2021 International Joint Conference on Neural Networks*, pp. 1-8, 2021, DOI: 10.1109/IJCNN52387.2021.9534060.
- [39] R.S. Sutton and A.G. Barto, “Reinforcement Learning: An Introduction”, *IEEE Transactions on Neural Networks*, vol. 9, no. 5, pp. 1054-1054, 1998, DOI: 10.1109/TNN.1998.712192.
- [40] TensorCell, the official group’s website, <https://www.tensorcell.com>. Last accessed: 12.06.2023.
- [41] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, N.A. Gomez, ŁKaiser, and I. Polosukhin, “Attention is All you Need”, *Advances in Neural Information Processing Systems*, vol. 30, pp. 5998-6008, 2017.
- [42] S. Wolfram, “A new kind of science”, Wolfram Media, ISBN: 1-57955-008-8, 2002.

- [43] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, et al., “A Comprehensive Survey on Transfer Learning”, *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43-76, 2021, DOI: 10.1109/JPROC.2020.3004555.