

Identyfikacja deterministycznych i stochastycznych automatów komórkowych na podstawie obserwacji cząstkowych

Witold Bołt

Warszawa, 2023

Niniejszy dokument stanowi **streszczenie rozprawy doktorskiej** napisanej pod kierunkiem promotora **prof. dr Bernarda De Baets** oraz promotora pomocniczego **dr hab. Jana W. Owińskiego**, w dziedzinie naukowej **informatyka techniczna i telekomunikacja**, złożonej do obrony w **Inście Badań Stosowanych Polskiej Akademii Nauk** w Warszawie. Rozprawa doktorska przedstawiona jest formie tematycznie spójnego cyklu publikacji naukowych i napisana jest w całości w **języku angielskim**.

1 Wstęp

Automaty Komórkowe (CAs, z *ang.* Cellular Automata) to układy dynamiczne, w których przestrzeń, czas i zbiór stanów są dyskretne (a często wręcz skończone). Pomimo tak prostej charakteryzacji, CAs wykazują wiele złożonych właściwości. Ze względu na te dwa czynniki - prostą definicję i interesujące właściwości - CAs są ważnym narzędziem do modelowania różnego rodzaju zjawisk np. fizycznych, a także istotnym obiektem badań teoretycznych.

Niniejsza praca doktorska poświęcona jest problemowi identyfikacji CAs, czyli automatycznemu znajdowaniu CA odpowiadającego danemu zbiorowi obserwacji. Innym słowy, algorytm identyfikacji przyjmuje zestaw, potencjalnie niekompletnych, „obrazów” przedstawiających ewolucję pewnego nieznanego CA i stara się znaleźć CA, który jest w stanie odtworzyć te same „obrazy”, uzupełniając brakujące informacje. Rozważane tu obserwacje, czy też mówiąc mniej formalnie „obrazy”, należy rozumieć jako abstrakcyjny pomiar rzeczywistości – zapis ewolucji pewnego nieznanego systemu, który ma być modelowany za pomocą CAs.

Zagadnienie identyfikacji rozważane jest dla dwóch klasy jednowymiarowych, binarnych CAs, tj. deterministycznych i niedeterministycznych CAs. Dla obu tych klas zaproponowano algorytm identyfikacji, pozwalający na skuteczne znalezienie CA na podstawie zbioru obserwacji, które mogą być niekompletne (zamiennie używamy również terminu *cząstkowe*), przy czym niekompletność może się wiązać zarówno z brakiem informacji o stanie konkretnych komórek w konkretnych krokach czasowych jak i brakiem zapisów całych kroków czasowych. Zakłada się przy tym, że dokładna liczba brakujących wierszy (kroków czasowych) jest również nieznana. Taki sposób ujęcia problemu motywowany jest praktycznymi aspektami modelowania zjawisk rzeczywistych.

Identyfikacja CAs omawiana jest tu w ujęciu teoretycznym, odnosząc się jedynie do wybranych klas jednowymiarowych, binarnych CAs. Dzięki temu możliwa była bardzo szczegółowa analiza wyników szerokiej gamy eksperymentów. Konstrukcja proponowanych algorytmów identyfikacji pozwala jednak na ich proste rozszerzenie do bardziej ogólnych przypadków (CAs w wielu wymiarach i/lub z wieloma stanami).

2 Artykuły wchodzące w skład rozprawy doktorskiej

Następujące artykuły wchodzą w skład prezentowanej rozprawy doktorskiej:

- [B1] W. Bołt, J. M. Baetens i B. De Baets, „An Evolutionary Approach to the Identification of Cellular Automata Based on Partial Observations,” w *2015 IEEE Congress on Evolutionary Computation (CEC)*, 2015, s. 2966–2972. DOI: 10.1109/CEC.2015.7257258.
- [B2] W. Bołt, B. Wolnik, J. M. Baetens i B. De Baets, „On the Identification of α -Asynchronous Cellular Automata in the Case of Partial Observations with Spatially Separated Gaps,” w *Challenging Problems and Solutions in Intelligent Systems*, G. d. Trè, P. Grzegorzewski, J. Kacprzyk, J. W. Owsiniński, W. Penczek i S. Zadrozny, red. Springer International Publishing, 2016, s. 23–36, ISBN: 978-3-319-30165-5. DOI: 10.1007/978-3-319-30165-5_2.
- [B3] W. Bołt, A. Bołt, B. Wolnik, J. M. Baetens i B. De Baets, „A Statistical Approach to the Identification of Diploid Cellular Automata,” w *Theory and Practice of Natural Computing*, C. Martín-Vide, R. Neruda i M. A. Vega-Rodríguez, red., Springer International Publishing, 2017, s. 37–48, ISBN: 978-3-319-71069-3. DOI: 10.1007/978-3-319-71069-3_3.

- [B4] W. Bołt, A. Bołt, B. Wolnik, J. M. Baetens i B. De Baets, „A Statistical Approach to the Identification of Diploid Cellular Automata Based on Incomplete Observations,” *Biosystems*, t. 186, s. 103–976, 2019. DOI: 10.1016/j.biosystems.2019.103976.
- [B5] W. Bołt, J. M. Baetens i B. De Baets, „Identification of Cellular Automata Based on Incomplete Observations With Bounded Time Gaps,” *IEEE Transactions on Cybernetics*, t. 50, nr. 3, s. 971–984, 2020. DOI: 10.1109/TCYB.2018.2875266.

Prezentowane artykuły stanowią spójny tematycznie cykl prac, który podzielić można na dwie części. W artykułach [B1], [B5] omówiono zagadnienie identyfikacji w przypadku deterministycznych CAs, natomiast w pracach [B2]–[B4] prezentowane są wyniki dla wybranych klas stochastycznych CAs.

3 Podsumowanie głównych wyników

3.1 Preliminaria

Deterministyczny CA to układ dynamiczny, w którym dyskretna przestrzeń podzielona jest na niepodzielne elementy zwane komórkami. W czasie (który również jest dyskretny), każda z komórek przyjmuje jeden ze stanów z ustalonego zbioru stanów, który z reguły jest skończony. Zmiana stanów komórek następuje zgodnie z ustaloną regułą i zależy od stanu danej komórki oraz od stanów komórek w jej ustalonym sąsiedztwie (z reguły sąsiedztwo to komórki w jakimś sensie najbliższe). Zapis stanów wszystkich komórek automatu w danym kroku czasowym nazywam konfiguracją tego automatu. Zapis stanów komórek z sąsiedztwa danej komórki nazywamy konfiguracją sąsiedztwa. Ewolucja automatu komórkowego zaczyna się od zadanej konfiguracji początkowej. W tej pracy rozważamy jedynie CAs, w których komórki tworzą skończoną przestrzeń jednowymiarową, którą można utożsamiać z podziałem okręgu na pewną liczbę elementów. Ponadto rozważamy tu jedynie CAs binarne, czyli takie, w których zbiór stanów jest postaci $\{0, 1\}$. Załóżmy, że N oznacza liczbę komórek, a $\mathbf{x}^t = (x_0^t, \dots, x_{N-1}^t)$ to konfiguracja automatu w chwili czasu t . Dla dowolnego $n = 0, \dots, N - 1$ stan n -tej komórki w chwili czasu $t + 1$, oznaczony jako x_n^{t+1} wylicza się za pomocą ustalonej reguły lokalnej $f: \{0, 1\}^{2r+1} \rightarrow \{0, 1\}$ zgodnie ze wzorem:

$$x_n^{t+1} = f(x_{n-r}^t, \dots, x_n^t, \dots, x_{n+r}^t), \quad (1)$$

przy czym wszystkie operacje na indeksach wykonywane są modulo N . Liczbę $r \geq 0$ określającą rozmiar sąsiedztwa nazywamy promieniem. Reguła lokalna

definiuje w naturalny sposób regułę globalną $F: \{0, 1\}^N \rightarrow \{0, 1\}^N$ następująco:

$$F(\mathbf{x}^t) = \mathbf{x}^{t+1} = (x_0^{t+1}, x_1^{t+1}, \dots, x_{N-1}^{t+1}). \quad (2)$$

Przyjmując konwencję zapisu wszystkich konfiguracji sąsiedztwa, których jest 2^{2r+1} , w porządku leksykograficznym, możemy jednoznacznie wyrazić regułę lokalną f jako wektor binarny $(\ell_0, \dots, \ell_{2^{2r+1}})$. Liczba ℓ_i oznacza wartość reguły f na i -tym sąsiedztwie w porządku. Binarne, deterministyczne, jednowymiarowe CAs, dla których istnieje reguła lokalna o promieniu $r = 1$ nazywamy Elementarnymi CAs (ECAs).

Skończony ciąg konfiguracji CA postaci $(\mathbf{x}^0, \mathbf{x}^1, \dots, \mathbf{x}^{T-1})$ nazywamy diagramem czaso-przestrzennym. Diagram taki można przedstawić w postaci macierzy, w której wiersze odpowiadają konfiguracjom w kolejnych krokach czasu. Częstkowe (niekompletne) zapisy diagramów czaso-przestrzennych używane są jako dane wejściowe w algorytmie identyfikacji CA omawianym w niniejszej pracy.

Rozszerzenie definicji CA podanej wyżej na przypadek Stochastycznych CA (SCA), sprowadza się do modyfikacji definicji reguły lokalnej i globalnej, które stają się funkcjami losowymi przypisującymi stany zgodnie z ustalonym rozkładem prawdopodobieństwa. Rozważamy jedynie przypadek, w którym:

$$x_n^{t+1} = X_{t,n}(x_{n-r}^t, \dots, x_n^t, \dots, x_{n+r}^t), \quad (3)$$

gdzie $X_{t,n}$ są niezależnymi zmiennymi losowymi o rozkładzie Bernoulliego spełniającymi:

$$\Pr(X_{t,n}(x_{n-r}, \dots, x_n, \dots, x_{n+r}) = 1) = p_i, \quad (4)$$

gdzie i odpowiada numerowi konfiguracji sąsiedztwa $(x_{n-r}, \dots, x_n, \dots, x_{n+r})$, czyli prawdopodobieństwo przejścia n -tej komórki w stan 1 (lub 0) zależy jedynie of konfiguracji sąsiedztwa, ale nie zależy czasu t ani od numeru komórki n . Wektor prawdopodobieństw (p_i) zapisany zgodnie z porządkiem konfiguracji sąsiedztwa pozwala jednoznacznie zdefiniować regułę lokalną SCA. Jest on naturalnie uogólnieniem wektora (ℓ_i) z przypadku deterministycznego CA.

Pojęcie diagramu czaso-przestrzennego dla SCA zyskuje szersze znaczenie. Dla ustalonej konfiguracji początkowej \mathbf{x}^0 diagram czaso-przestrzenny SCA to rozkład prawdopodobieństwa nad przestrzenią wszystkich ciągów konfiguracji. Szczegółowa charakteryzacja tego rozkładu jest trudna i mało przydatna w praktyce. Symulacje komputerowe SCA, pozwalają nam jednak próbkować ten rozkład poprzez generowanie jego konkretnych realizacji. Każda z nich jest diagramem czaso-przestrzennym w sensie deterministycznym – zapisem

jednej konkretnej trajektorii w przestrzeni konfiguracji. Właśnie takie realizacje stanowią dane wejściowe do algorytmów identyfikacji SCAs omawianych w niniejszej pracy.

Każdy SCA może zostać przedstawiony jako skończony wektor prawdopodobieństw (λ_k) sumujący się do jedności i skończony wektor deterministycznych CA (A_k), tak, że w każdej chwili czasu, dla każdej komórki niezależnie, stosowany jest losowo wybrany automat A_k , przy czym prawdopodobieństwo wyboru A_k wynosi λ_k . Takie przedstawienie nazywamy stochastyczną mieszaną CAs.

W tej rozprawie rozważamy szczególne przypadki stochastycznych mieszanek – diploidy i α -asynchroniczne CAs (α -ACAs). Diploid to SCA zdefiniowany przez stochastyczną mieszaną dwóch CAs A_1 i A_2 oraz prawdopodobieństwo λ . W każdym kroku czasu, dla każdej z komórek niezależnie, wybierany jest losowo jeden z dwóch CAs. Wybór A_1 następuje z prawdopodobieństwem λ , a A_2 z prawdopodobieństwem $1 - \lambda$. Natomiast α -ACAs, omawiane osobno ze względu na istotne znaczenie praktyczne, to szczególne diploidy, w których CA A_2 jest zadany przez regułę „identyczność” (stany komórek nie ulegają zmianom w czasie).

3.2 Identyfikacja deterministycznych CAs

W identyfikacji deterministycznych CAs stosujemy specjalnie zaprojektowany algorytm ewolucyjny będący wariantem dobrze znanego algorytmu genetycznego. Algorytm ten rozpoczyna się od populacji początkowej wybranych losowo potencjalnych rozwiązań, reprezentowanych jako binarne wektory (ℓ_i) kodujące reguły lokalne CA. Populacja początkowa jest modyfikowana za pomocą operatorów genetycznych tworząc nową populację. Następnie proces powtarza się. W ramach danej populacji dopuszcza się tu wektory różnej długości, dzięki czemu algorytm może badać CAs, których reguły lokalnej mają różny rozmiar sąsiedztwa. Wymaga to wprowadzenia odpowiedniej konstrukcji operatorów genetycznych, które potrafią działać na osobnikach różnej długości zachowując poprawne znaczenie poszczególnych bitów w ich zapisie.

Kluczowym elementem omawianego algorytmu, poza definicją operatorów genetycznych, jest funkcja dopasowania. W toku badań zostało rozważonych wiele wariantów tej funkcji. Jej najprostsza i w pewnym sensie „oczywista” postać porównująca całą obserwację, z całym obrazem czaso-przestrzennym badanego kandydata, okazała się nieskuteczna. Zamiast tego stosuje się podejście, w którym obserwacja traktowana jest jako zbiór niezależnych par, zbudowanych z kolejnych jej wierszy. Załóżmy, że I jest obserwacją. Rozważamy wówczas pary $I[1] \rightarrow I[2]$, $I[2] \rightarrow I[3]$ itd. Pierwszy z elementów pary $I[j]$ staje się ustaloną konfiguracją początkową CA. Startując z tej kon-

figuracji generowany jest diagram czaso-przestrzenny zawierający T kroków czasowych. Spośród nich wybierany jest taka konfiguracja, w której liczba różnic z konfiguracją $I[j + 1]$ jest najmniejsza. Ta liczba różnic staje się częścią składową błędu. Błędy policzone osobno dla każdej z par są sumowane i stanowią łączny błąd dopasowania kandydata do obserwacji. Algorytm poszukuje takiego kandydata, dla którego błąd wynosi 0.

Zgodnie z informacją zawartą we Wstępie, zakłada się, że obserwacje mogą być cząstkowe, tj. zarówno poszczególne stany konkretnych komórek mogą być nieznane (zamiast wartości 0 lub 1 zapisana jest wówczas wartość ?), jak i pewne konfiguracje (kroki czasowe) mogą być niezabserwowane. Procedura obliczająca błąd, pomija komórki oznaczone symbolem ?, a następnie wypełnia je wartościami wynikającymi z działania badanej reguły na potrzeby dalszych obliczeń. Wylizanie T kroków pomiędzy każdą z par konfiguracji odczytanych z obserwacji pozwala natomiast poradzić sobie z brakującymi „pośrednimi” konfiguracjami, które być może nie zostały zaobserwowane.

Wyniki eksperymentów obliczeniowych potwierdzają, że omówiony algorytm rozwiązuje problem nawet w przypadkach cząstkowych obserwacji z wieloma brakującymi komórkami i krokami czasowymi. Analiza otrzymanych wyników pokazuje silną zależność między właściwościami dynamicznymi identyfikowanego CA a wysiłkiem potrzebnym do jego identyfikacji (mierzonym liczbą iteracji algorytmu ewolucyjnego).

Szczegółowy opis algorytmu oraz opracowanie wyników eksperymentów znajduje się w pracach [B1], [B5]. W pracy [B1] przeanalizowano szczegółowo eksperymenty z identyfikacją dwóch wybranych ECAs, natomiast w [B5] przeanalizowano szerzej wszystkie ECAs oraz szersze klasy CAs o większych sąsiedztwach.

3.3 Identyfikacja niedeterministycznych CAs

Identyfikacja SCA ograniczona jest do przypadku dwóch klas automatów: α -ACAs oraz diploidów. Zakłada się tu, że znany jest promień sąsiedztwa reguł lokalnych i rodzaj poszukiwanego SCA – tj. z góry wiadomo czy jest to α -ACA czy diploid.

Do identyfikacji SCA stosowany jest deterministyczny algorytm konstrukcji rozwiązania, oparty na wnioskowaniu statystycznym i estymacji parametrów. Na podstawie obserwacji wylizane są tablice częstości wystąpień dla konfiguracji sąsiedztw – tablica L , a także częstości przejścia centralnej komórki w stan 1 w następnym kroku czasu – tablica K . Sąsiedztwa w których występuje niekompletność obserwacji oznaczona przez symbol ? są pomijane w zliczaniu. Dzięki temu, że możemy mieć wpływ na konfigurację początkową,

zakładamy, że $L_i \neq 0$ dla każdego i , to znaczy każda konfiguracja sąsiedztwa wystąpiła co najmniej raz w obserwacji. Łatwo pokazać wtedy, że iloraz $\frac{K_i}{L_i}$ jest estymacją prawdopodobieństwa p_i . Zależnie od obserwacji, dokładność takiej estymacji może być jednak bardzo niska. Korzystając z własności diploidów, można poprawić dokładność estymacji sumując niektóre z wyliczonych częstości używając je do estymacji prawdopodobieństwa λ . Stosowana jest standardowa metoda przybliżenia rozkładu Bernoulliego rozkładem normalnym i po weryfikacji kilku technicznych założeń otrzymywany jest przedział ufności dla λ a następnie poprzez analizując tablice K i L , budowane są reguły lokalne CAs A_1 i A_2 . Konstrukcja ta jest relatywnie prosta, jednak zapewnienie, że prawdopodobieństwo popełnienia błędu mieści się w ustalonym przedziale ufności wymaga dłuższej, technicznej dyskusji, która została zaprezentowana w [B3], [B4].

Po skonstruowaniu reguł CAs A_1 i A_2 oraz estymacji prawdopodobieństwa λ , algorytm dodatkowo pozwala na oszacowanie najbardziej prawdopodobnych wartości niezaobserwowanych stanów, czyli komórek oznaczonych w obserwacji przez ?. Analizując stany komórek w poprzedzającym i następnym wierszu obserwacji i rozważając wszystkie możliwości, proponowane jest wypełnienie o najwyższej wiarygodności, które gwarantuje zgodność z regułami CAs A_1 i A_2 . Innymi słowy daje gwarancję, że po wypełnieniu wszystkich luk ponowne wykonanie procedury identyfikacji wyłoni te same CAs A_1 i A_2 . Najbardziej naiwne podejście do wypełniania luk nie gwarantuje spełnienia tej własności. Stosowana jest nieco bardziej skomplikowana strategia, której szczegółowy opis przedstawiony jest w [B4].

Podobnie jak w przypadku deterministycznych CAs, efektywność proponowanego algorytmu zweryfikowano za pomocą licznych eksperymentów obliczeniowych. Przeprowadzone badania potwierdzają dużą skuteczność algorytmu, a także wskazują na zależności między właściwościami dynamicznymi identyfikowanych CAs oraz dokładnością identyfikacji.

Szczegółowy opis algorytmu oraz wyniki eksperymentów dla α -ACA znajdują się w [B2]. Natomiast w [B3], [B4] znajduje się szczegółowy opis algorytmu i analiza wyników licznych eksperymentów dla diploidów. Wszystkie prezentowane eksperymenty dotyczą SCAs zbudowanych w oparciu o ECAs, jednak metoda daje się łatwo uogólnić również na sąsiedztwa o większym promieniu.

4 Kierunki przyszłych prac naukowych

4.1 Przypadek obserwacji z szumem

W omawianym w tej pracy problemie identyfikacji, brane pod uwagę są jedynie obserwacje wolne od szumu. Co prawda niekompletność obserwacji, która jest w pewnym sensie losowa, może być utożsamiana ze specyficznym rodzajem szumu, to jednak praktyczne zastosowania CAs wymagają rozpatrzenia bardziej ogólnego przypadku. Chodzi mianowicie o przypadek, gdy nie ma pewności co do poprawności odczytania stanów pewnych komórek w procesie rejestrowania obserwacji. Innymi słowy, dla pewnych komórek zarejestrowano stan inny od faktycznego (w przypadku binarnym oznacza to oczywiście zamianę 0 z 1 lub 1 z 0). Autor opracował uogólnienie algorytmu identyfikacji deterministycznych CAs w oparciu o zaszumione, cząstkowe obserwacje. Uogólnienie to polega na modyfikacji procedury wyliczania funkcji dopasowania. W momencie wyliczania błędu wybranego kandydata na wybranej obserwacji, przydzielany jest ustalony parametrem „budżet poprawek”. Porównując obserwację z otrzymanym obrazem kandydata, dopuszczamy określoną tym budżetem liczbę rozbieżności, które nie powiększają błędu. Jeśli dana rozbieżność mieści się w budżecie, to nie dość, że nie powiększa błędu, to skutkuje również tymczasowym poprawieniem stanu w obserwacji. Ze względu na zależności pomiędzy sąsiednimi komórkami, a także między sąsiednimi krokami czasowymi, dokonanie takiej pojedynczej poprawki może mieć wpływ na wiele innych komórek i w efekcie istotny wpływ na otrzymaną wartość funkcji dopasowania (silniejszy niż tylko bezpośrednio pominięte różnice). Uogólnienie te zostało przebadane na ograniczonym zbiorze obserwacji, ale otrzymane wyniki są bardzo obiecujące. Przy założeniu, że budżet poprawek jest poprawnie ustalony, w oparciu o oszacowanie liczby komórek dotkniętych przez szum, algorytm jest w stanie poprawnie zidentyfikować CA, nawet gdy zmienionych jest kilkanaście procent komórek w obserwacji. Co więcej, w naturalny sposób identyfikacja prowadzi również do usunięcia szumu z obserwacji. Dalsze prace w tym kierunku zmierzają do przebadania algorytmu na większej liczbie przypadków.

4.2 Identyfikacja afinicznych, ciągłych CAs

Afiniczne, ciągłe CAs (ACCAs) to często używane uogólnienie binarnych CAs, polegające na zastąpieniu zbioru stanów $\{0, 1\}$ przedziałem $[0, 1]$, a reguł lokalnych pewną specjalną formą wielomianów o współczynnikach z przedziału $[0, 1]$. Wielomiany te wywodzą się z tzw. *fuzyfikacji* (inaczej, rozmycia – w sensie logiki rozmytej) reguł binarnych CAs przedstawionych jako wyrażenia

logiczne. Ponieważ ACCAs to bardzo naturalne uogólnienie binarnych CAs, możliwe jest podobnie naturalne uogólnienie algorytmu identyfikacji na przypadek ACCAs. W miejsce prezentowanego algorytmu ewolucyjnego opartego na algorytmie genetycznym stosować można algorytm ewolucji różnicowej. Funkcja dopasowania w tym algorytmie jest niemal identyczna jak ta omówiona dla binarnych CAs, z tą różnicą że zamiast liczyć błąd w oparciu o liczbę komórek z niepasującymi stanami, liczy się kwadrat różnicy między stanami w obserwacji a stanami otrzymanymi z badanego kandydata na rozwiązanie. Autor przedstawił wstępne, pozytywne wyniki odnośnie tego uogólnienia algorytmu na konferencji Summer Solstice 2015 w Toronto. Dalsze prace są prowadzone aby pokazać efektywność algorytmu w zależności o własności dynamicznych identyfikowanych ACCAs.

4.3 Identyfikacja stochastycznych mieszanek CAs

W przedstawionej pracy omówiono algorytm identyfikacji SCAs w przypadku wybranych klas tych automatów, tj. diploidów i α -ACAs. Każdy SCA może zostać przedstawiony jako stochastyczna mieszanka skończonej liczby CAs. Przedstawiony algorytm identyfikacji, choć zastosowany tu jedynie dla stochastycznych mieszanek dwóch CAs, daje się jednak relatywnie prosto uogólnić dla większej liczby składników. Autor prowadził wstępne badania i symulacje dla mieszanek trzyskładnikowych. Otrzymane wyniki są bardzo obiecujące. Dalsze badania w tym kierunku mają na celu przeanalizowanie uogólnienia algorytmu na stochastyczne mieszanki dowolnie wielu CAs.

4.4 Metody uczenia maszynowego

Wielu autorów w ostatnich latach przedstawiło wyniki badań łączących popularne techniki uczenia maszynowego z modelami opartymi o CAs. W szczególności pokazano, że splotowe sieci neuronowe (*ang.* Convolutional Neural Networks – CNNs) potrafią symulować pewne klasy CAs. Zbadano również modele hybrydowe, w których rolę lokalnej reguły CA pełni szczególny rodzaj sieci neuronowej. Wyniki tych badań są bardzo obiecujące i motywują próby powiązania algorytmów identyfikacji z sieciami neuronowymi.

Rozważane są dwa podejścia. Pierwsze, w którym sieć neuronowa używana jest do przyspieszenia działania i poprawy efektywności algorytmu ewolucyjnego, poprzez aproksymację funkcji dopasowania. W trakcie ewolucji algorytmu identyfikacyjnego, zbierane są wszystkie uzyskane wyniki funkcji dopasowania i gdy zarejestruje się ich odpowiednio dużo, trenuje się sieć, której celem jest przybliżanie wartości tej funkcji. W kolejnych krokach ta

aproksymacja używana jest zamiast właściwej funkcji, co znacznie przyspiesza prace algorytmu ewolucyjnego, w którym główny koszt obliczeniowy stanowi właśnie wyliczanie wartości funkcji dopasowania.

Drugie z podejść to budowa uniwersalnego klasyfikatora na potrzeby identyfikacji zadanej klasy CAs, wytrenowanego na bardzo dużym zbiorze obserwacji wygenerowanych przez różne CAs. Innymi słowy, w podejściu tym problem identyfikacji CAs sprowadzony jest do problemu uczenia nadzorowanego (*ang.* supervised learning), w którym dane wejściowe to obserwacje a ich etykiety, użyte w procesie uczenia, to wektory (ℓ_i) reprezentujące reguły lokalne.

Oba podejścia są obecnie wstępnie badane przez autora i otrzymane dotychczas wyniki pierwszych symulacji są bardzo obiecujące.

5 Podsumowanie

W ramach powyższej rozprawy omówiono i rozwiązano problem identyfikacji CAs, który inspirowany jest praktycznymi zastosowaniami CAs. Praca ta stanowi krok w kierunku automatycznego budowania modeli CA w oparciu o dane z eksperymentów lub zapisu zjawisk rzeczywistych. Dalsze prace prowadzone obecnie przez autora, w kierunkach wskazanych w Rozdziale 4, zmierzają do udoskonalenia prezentowanych metod i udostępnienia ich szerszemu gronu badaczy.