

Mirosława Lasek

Uniwersytet Warszawski
e-mail: mlasek@wne.uw.edu.pl

Witold Lasek

Warszawski Uniwersytet Medyczny
e-mail: witold.lasek@wum.edu.pl

Marek Pęczkowski

Uniwersytet Warszawski
e-mail: mpeczkowski@wne.uw.edu.pl

OD IMMUNOLOGII DO MODELOWANIA, PRZETWARZANIA I ANALIZ DANYCH

Streszczenie: Podglądanie przyrody jest często inspiracją dla niekonwencjonalnych i nowatorskich pomysłów. Sztuczne sieci neuronowe, algorytmy genetyczne i ewolucyjne powstały jako naśladowanie rozwiązań biologicznych. W ostatnich latach następuje szybki rozwój algorytmów immunologicznych. Algorytmy te znajdują zastosowanie w matematyce, technice, zarządzaniu, ekonomii, finansach. Na początku artykułu opisujemy działanie układu immunologicznego stanowiącego podstawę budowy algorytmów immunologicznych. Następnie przedstawiamy mechanizmy kształtujące algorytmy układu immunologicznego, w szczególności negatywnej selekcji, sieci immunologicznej, selekcji klonalnej. Wskazujemy w jaki sposób elementy i mechanizmy sztucznego układu immunologicznego są reprezentowane w algorytmach umożliwiającym modelowanie, przetwarzanie i analizy danych. W ostatniej części artykułu, wymieniamy przykłady zastosowań.

Słowa kluczowe: immunologia, immunocomputing, sztuczny układ immunologiczny (artificial immune system), algorytmy sztucznych układów immunologicznych.

1. Wstęp

Mechanizmy działania biologicznego układu immunologicznego (odpornościowego), takie jak negatywna selekcja, działanie sieci immunologicznych ujmujących zjawiska interakcji idiotypowych i antyidiotypowych, selekcja klonalna, zagrożenia w odpowiedzi immunologicznej [Gołąb i in. (red) 2012; Lasek 2009], stały się inspiracją dla rozwoju modelowania, przetwarzania i analiz danych, rozpatrywanych

z dwóch punktów widzenia, prezentujących nieco różniące się od siebie ujęcia i podejścia [De Castro 2006, s. 267]. Pierwszy punkt widzenia skupia się na obliczeniach i przetwarzaniu komputerowym (*immunocomputing*), rozumianym jako nowe podejście, ukierunkowane na dążenie do wykorzystania sposobów obróbki informacji wywodzących się od mechanizmów właściwych dla biologicznego układu immunologicznego i prowadzących do komputerowych algorytmów i oprogramowania, aby ostatecznie osiągnąć koncepcję, w której komputer działa jako urządzenie przetwarzające w oparciu na zasadach biologicznego układu immunologicznego (*immunocomputer*). Drugi ujęcie to dążenie do opracowania sztucznego układu immunologicznego (*artificial immune system (AIS)*), zainspirowanego dorobkiem teoretycznym i eksperymentalnym immunologii, w którym celem jest opracowanie sposobów rozwiązywania problemów, np. ekonomicznych, podejmowania decyzji, zarządzania. Dąży się tu do wykorzystania różnorodnych mechanizmów zaczerpniętych z metod rozwiązywania konkretnych problemów przez rzeczywisty układ immunologiczny. Zostały już opracowane odpowiedniki pewnych pojęć i koncepcji rzeczywistego układu immunologicznego, które nawet po ich zaledwie pobieżnym przeanalizowaniu wydają się trafną analogią, mogącą zapewnić udane naśladownictwo biologicznego pierwowzoru i być źródłem nowych pomysłów dla przetwarzania komputerowego i rozwiązywania problemów, a wywodzących się mniej lub bardziej pośrednio „z metod naturalnych, jakie wykształciła przyroda”. W pracy [De Castro 2006, s. 312] wybrane pojęcia i koncepcje biologicznego układu immunologicznego oraz ich odpowiedniki, które wydają się adekwatne dla utworzenia układu sztucznego, zestawiono w postaci tabelarycznej. Podstawowe pojęcia i koncepcje zostaną omówione w niniejszym artykule. Głównymi cechami biologicznych układów immunologicznych, skłaniającymi do uznania, że mechanizmy działania tych układów mogą być inspiracją do tworzenia oryginalnych rozwiązań w dziedzinie przetwarzania komputerowego i opracowywania sposobów rozwiązywania problemów są m.in.: brak potrzeby scentralizowanego nadzoru działania elementów, możliwość wykonywania obliczeń w sposób równoległy, elastyczność tworzenia rozwiązań i dostosowanie metody działania do wymogów problemu, możliwość wykrywania anomalii i elementów „obcych”, nie napotkanych uprzednio, adaptacyjność systemu, pozwalająca na jego przystosowanie się do zmieniających się warunków, łatwość przedstawiania struktur i mechanizmów systemu w sposób sformalizowany (por. też [Wierchoń 2001, s. 1-2; Dasgupta, Niño 2009, s. 40-41; Kołodziejczyk]).

Celem artykułu jest przedstawienie możliwości wykorzystania mechanizmów biologicznego układu immunologicznego w modelowaniu, przetwarzaniu i analizach danych.

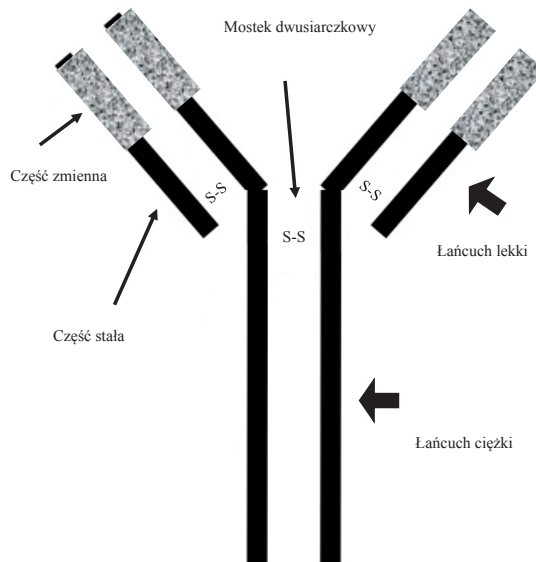
Dla czytelników, którym nieznaną jest dziedzina immunologii, krótko opisano układ immunologiczny żywych organizmów, zwracając uwagę na takie podstawowe elementy ważne w działaniu tego układu, jak narządy limfatyczne i fizjologia układu oraz mechanizm rozpoznawania antygenów przez limfocyty. W celu pełniejszego zapoznania czytelnika z działaniem układu immunologicznego omówiono bardziej

szczegółowo niezwykle istotne dla sposobu jego funkcjonowania dwie podstawowe teorie działania układu: teorię sieci immunologicznej, uwzględniającej oddziaływania typu „przeciwciało – antygen” i typu „przeciwciało – przeciwciało”, a także teorię zagrożenia. W dalszej części artykułu zajęto się problemem, jak zbudować elementy systemów: komputerowego przetwarzania danych i rozwiązywania problemów (systemów sztucznych), które stanowiłyby elementy ich biologicznych odpowiedników (systemu rzeczywistego). Podjęto tu dwa podstawowe, kluczowe zagadnienia: jak odwzorować w modelu działanie biologicznych mechanizmów: jakie rodzaje algorytmów wybrać i w jaki sposób je konstruować, aby jak najlepiej odzwierciedlały realizację mechanizmów immunologicznych, oraz jak reprezentować elementy składowe systemów, będące odpowiednikami biologicznych tworów (np. komórek, molekuł) oraz biologicznych zasad funkcjonowania układu immunologicznego. Ostatnią część artykułu poświęcono zilustrowaniu, za pomocą przykładów, możliwości zastosowań mechanizmów układu immunologicznego w modelowaniu, przetwarzaniu i analizach danych. Wspomniano o rozwiązaniach już zrealizowanych zastosowań zainspirowanych biologicznymi mechanizmami, m.in.: o negatywnej selekcji, selekcji klonalnej, sieci immunologicznych. Mechanizmy te należą do mechanizmów zachowań biologicznego układu immunologicznego o ugruntowanych, mocnych podstawach teoretycznych, dobrze uzasadnionych i zweryfikowanych doświadczalnie i klinicznie. Zastosowane w całkiem innej dziedzinie niż „biologiczny pierwowzór”, pozwalają łatwo dostrzec inspirację z tego „biologicznego pierwowzoru” i skłaniają do zastanowienia się nad korzyściami przystosowania rozwiązań z odległych i nie mających wiele wspólnego dziedzin, jak biologia i modelowanie, przetwarzanie i analiza danych.

2. Układ immunologiczny żywych organizmów

Układ immunologiczny człowieka, podobnie jak innych zwierząt, opisany w punkcie 2. artykułu na podstawie szczegółowo przedstawiających go prac: [Gołąb i in. (red.) 2012; Lasek 2009], pełni funkcje obronne względem czynników zakaźnych, określanych jako patogeny (*pathogens*). Należą do nich wirusy, wiele bakterii, niektóre gatunki grzybów i pierwotniaków, jak również wielokomórkowe pasożyty, np. nicienie i tasiemce. W połączeniu z innymi układami człowieka, nerwowym i układem gruczołów wydzielania wewnętrznego (endokrynowym), układ immunologiczny umożliwia przeżycie i broni integracji organizmu. Współdziała on z barierami chemicznymi i fizycznymi, chroniącymi przed zakażeniami skórę i błony śluzowe. Należą do nich m.in. śluz z zawartymi czynnikami antymikrobowymi, niskie pH skóry, substancje obronne w pocie i łoju itp. W obrębie układu immunologicznego działają dwa główne mechanizmy obronne: naturalne (wrodzone; *innate*) i nabyte (*adaptive*). Te pierwsze warunkują nieswoistą, te drugie – swoistą odpowiedź immunologiczną. W odpowiedzi immunologicznej nieswoistej uczestniczą komórki układu immunologicznego, takie jak makrofagi i granulocyty obojętnochłonne, jak rów-

niez wiele rozpuszczalnych cząsteczek obronnych. Odpowiedź ta stanowi pierwszą linię obrony przed patogenami. Jej najważniejszą cechą jest szybkość wystąpienia reakcji. Jest jednocześnie mało precyzyjna – nakierowana na charakterystyczne dla patogenów cząsteczki docelowe. Za swoistą obronę odpowiedzialne są limfocyty należące do krwinek białych – najważniejsze komórki układu immunologicznego. W grupie tych komórek mieszczą się dwie główne populacje: limfocyty *T* (*T lymphocytes, T cells*), powstające w grasicy (*thymus*, stąd nazwa tych komórek) i limfocyty *B* (*B lymphocytes, B cells*), rozwijające się w szpiku (*bone marrow*). Limfocyty *T* odpowiedzialne są za odpowiedź typu komórkowego, podczas której dochodzi do bezpośredniego zabijania zagrażających komórek, np. komórek zakażonych wirusami lub rozsiewającymi wirusy. Rolę tę pełnią limfocyty *T* cytotoksyczne (*Tc, cytotoxic T cells*). Limfocyty *T* mają również funkcje pomocnicze, takie jak koordynacja i profilowanie odpowiedzi immunologicznej. Wynika to ze zdolności wydzielania przez te limfocyty, określane jako limfocyty *Th* (*Th, T helper cells*), rozpuszczalnych cząsteczek (cytokin) oddziałujących na różne komórki. Limfocyty *B* i ich komórki potomne, określane jako komórki plazmatyczne (plazmocyty), uczestniczą w odpowiedzi immunologicznej humoralnej (łac. *humor* – płyn), w której za efekt obronny odpowiedzialne są działające swoiście rozpuszczalne cząsteczki nazywane przeciwciałami (immunoglobulinami) (*antibodies, immunoglobulins*). Przeciwciała składają się z czterech łańcuchów białkowych (dwóch identycznych ciężkich i dwóch takich samych lekkich), w każdym z nich można wyróżnić część zmienną i część stałą (rys. 1). Należy podkreślić, że odpowiedź immunologiczna swoista rozwija się wol-



Rys. 1. Budowa przeciwciała

Źródło: opracowanie własne.

niej niż nieswoista, jest jednak bardziej precyzyjna – rozwija się w stosunku do unikatowych cząsteczek na powierzchni patogenów, w tym bakterii, określanych jako antygeny (*antigens*). Gwoli ścisłości należy dodać, że w sensie immunologicznym antygenem jest każda cząsteczka (w wyjątkowych sytuacjach pochodząca z własnych komórek i tkanek), którą są w stanie rozpoznać komórki układu immunologicznego. Inną bardzo ważną cechą odpowiedzi immunologicznej swoistej jest tzw. pamięć immunologiczna, w wyniku której z reguły nie dochodzi do powtórnego zakażenia tym samym mikroorganizmem. Zjawisko to uzależnione jest od wykształcenia komórek pamięci (*memory cells*) wywodzących się zarówno z limfocytów *T*, jak i *B*, a jego przykładem jest nabywanie odporności na określone wirusy i bakterie po przechorowaniu określonej infekcji w dzieciństwie lub po szczepieniu profilaktycznym.

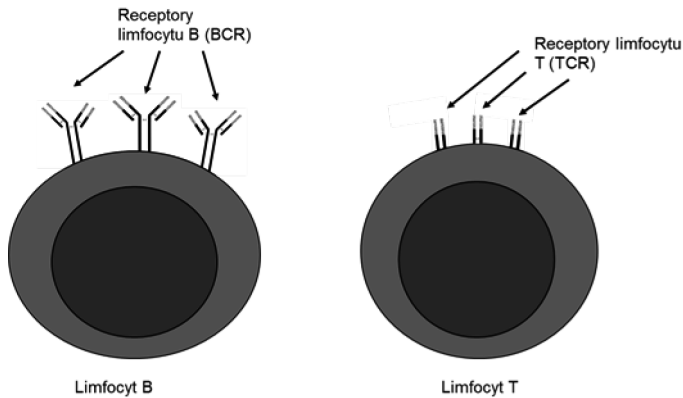
2.1. Narządy limfatyczne i działanie (fizjologia) układu immunologicznego

Jak wspomniano, limfocyty *T* i *B* (należące do krwinek białych) dojrzewają, odpowiednio, w grasicy i szpiku. Narządy te nazywane są centralnymi narządami limfatycznymi, choć szpik jest dodatkowo narządem krwiotwórczym, w którym powstają erytrocyty (krwinki czerwone) i inne krwinki białe (granulocyty i monocyty). W pełni sprawne limfocyty przedostają się do krwi i są rozprowadzane do obwodowych narządów limfatycznych, gdzie przebywają czasowo (przeważnie kilkanaście godzin), po czym odprowadzane są z powrotem, zazwyczaj z limfą, do krwi. Do obwodowych narządów limfatycznych należą węzły limfatyczne, migdałki, grudki limfatyczne skupione (kępki Peyera) oraz śledziona. Wtórne narządy limfatyczne są miejscem nagromadzenia dużej ilości limfocytów *B* i *T*. Dochodzi w nich do indukcji odpowiedzi immunologicznej. Limfocyty *B* rozpoznają antygeny, które dostały się do narządu limfatycznego w formie rozpuszczalnej (w przypadku węzłów limfatycznych – z limfą, w przypadku śledziony – z krwią), natomiast limfocyty *T* rozpoznają antygeny uprzednio przechwycone i przygotowane przez tak zwane komórki prezentujące antygen (*APC*, *antigen presenting cells*), z których najważniejsze to komórki dendryczne. Komórki te rozmieszczone są w skórze i błonach śluzowych i każdorazowo – po pochłonięciu mikroorganizmu lub jego fragmentu – wędrują do obwodowych narządów limfatycznych. W narządach tych, oprócz rozpoznania antygenów przez limfocyty *B* i *T*, dochodzi do kooperacji tych komórek, do pełnej aktywacji limfocytów *B* i produkcji przeciwciał niezbędna jest bowiem pomoc limfocytów *T* – wspomnianych wcześniej limfocytów pomocniczych *Th*. Wśród limfocytów *T* są również komórki wykazujące działanie hamujące reakcje immunologiczne – tzw. limfocyty *T* regulatorowe (*Treg*, *regulatory T cells*), niezbędne na przykład do tłumienia odpowiedzi na antygeny własne organizmu.

2.2. Rozpoznawanie antygenów przez limfocyty

Limfocyty *B* mają na powierzchni receptory, tj. cząsteczki identyfikujące antygeny, określane jako *BCR* (*B cell receptors*) – rys. 2, które mają budowę analogiczną do

przeciwciał i rozpoznają antygeny krążące w formie rozpuszczalnej. Natomiast limfocyty *T* mają receptory *TCR* (*T cell receptors*) (rys. 2), które z reguły rozpoznają antygeny peptydowe (białkowe) powstałe z odpowiednio pociętych w komórkach dużych cząsteczek. Antygeny te prezentowane są przez cząsteczki głównego układu zgodności tkankowej (*MHC*, *major histocompatibility complex*). Zarówno *BCR*, jak i *TCR* mają w budowie części zmienne, unikatowe dla każdego limfocytu, i części stałe, identyczne dla wszystkich limfocytów. Części zmienne receptorów są tymi miejscami, które zaangażowane są bezpośrednio w interakcje z antygenami.



Rys. 2. Limfocyt *B* i *T* oraz ich receptory rozpoznające antygen, odpowiednio *BCR* i *TCR*

Źródło: opracowanie własne.

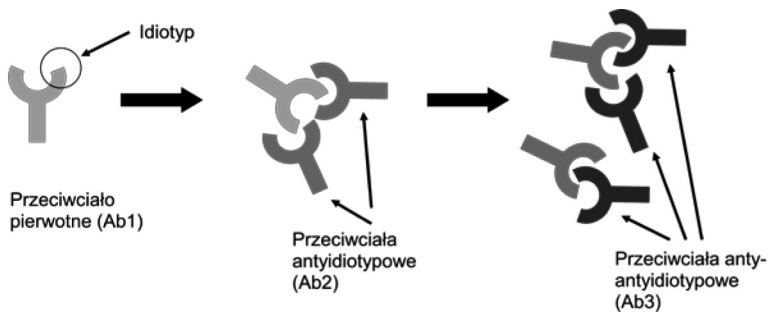
Ponieważ organizm człowieka styka się z milionami różnych antygenów, powstaje pytanie: skąd się bierze olbrzymia różnorodność receptorów, adekwatna do liczby rozpoznawanych antygenów? Okazało się, że w przeciwieństwie do praktycznie wszystkich białek organizmu łańcuchy białkowe receptorów *BCR* i *TCR* nie są kodowane przez gotowe geny, lecz powstają na podstawie powstałego genu tworzonego z dowolnie łączonych segmentów genowych w procesie rekombinacji w trakcie dojrzewania limfocytów. Segmenty te należą do tzw. bibliotek genów: *V*, *D* i *J* (dla części zmiennych receptorów) i *C* (dla części stałych). W obrębie biblioteki *V*, przykładowo dla receptora *BCR*, a konkretnie łańcucha ciężkiego tego receptora, u człowieka mieści się 51 segmentów o identycznej wielkości, lecz różniących się nieznacznie w strukturze, w bibliotece *D* – 27 segmentów, a w bibliotece *J* – 6 funkcjonalnych segmentów. Segmenty *V*, *D* i *J*, uogólniając, łączą się w sposób dowolny (w procesie rekombinacji), dając w efekcie olbrzymią liczbę kombinacji. Dodatkowo różnorodność powstałych ostatecznych genów jest zwiokrotniana dzięki zjawisku mutacji (przypadkowych zmian pojedynczych elementów genów – nukleotydów) w obrębie każdego segmentu, zachodzącemu podczas podziału komórek. Mutacja może doprowadzić do powstania ostatecznie receptora różniącego się nieznacznie

od oryginału. Receptor ten może wiązać się skuteczniej (z większą siłą, określoną powinowactwem do antygeny (*affinity*)), niż receptor powstały z genu pierwotnego. Jest to o tyle istotne, że w trakcie zakażenia bakteriami, podczas wielokrotnych podziałów tych patogenów, ich geny też mutują, dając warianty antygenów nieco zmienione i – jednocześnie – gorzej rozpoznawane przez limfocyty z wariantami receptora pierwotnego. Opisane zjawiska są fundamentem teorii selekcji klonalnej (*clonal selection*). Najważniejsze z postulatów tej teorii, to: (i) każdy limfocyt ma jeden rodzaj receptorów o unikatowej swoistości, (ii) rozpoznanie antygeny przez receptory limfocyty wiąże się z aktywacją komórki i powstaniem wielu komórek potomnych – klonów (*clones*) o identycznej swoistości, (iii) limfocyty mające receptory swoiste w stosunku do własnych antygenów giną w procesie selekcji rozwojowej we wczesnym etapie. Teoria selekcji klonalnej tłumaczy zdolność do reakcji i zwalczania praktycznie każdego czynnika zakaźnego (drobnoustroju) przez zainfekowany organizm. Uznaje się, że z olbrzymiej liczby limfocytów, których swoiste receptory zostały wygenerowane w procesie rekombinacji genów, w kontakcie z antygenami drobnoustroju szybko mnożą się wyselekcjonowane klony limfocytów, które najlepiej i najdokładniej rozpoznają dany antygen. Po wykształceniu swoistej odporności na patogen i po jego wyeliminowaniu dochodzi do wygaszenia odpowiedzi immunologicznej i przeżycia tylko części limfocytów – warunkujących pamięć immunologiczną. Po ponownym kontakcie z patogenem, dzięki komórkom pamięci, dochodzi do wtórnej odpowiedzi immunologicznej, która ma przebieg szybszy i znacznie intensywniejszy od odpowiedzi pierwotnej. W tym samym czasie pierwszorazowy kontakt z innym antygenem ma charakter odpowiedzi pierwotnej, porównywalnej co do nasilenia i czasu trwania do odpowiedzi na wcześniejsze antygeny. Powyższe rozważania odnoszą się do antygenów znacznie różniących się pod względem budowy. Zdarza się jednak, że w przypadku nieznacznie różniących się antygenów, dzięki zjawisku reaktywności krzyżowej (*ancross-reactivity*), organizm jest „częściowo przygotowany” do obrony i odpowiedź pierwotna zachodzi szybciej niż zazwyczaj. Swoistą, adaptacyjną odpowiedź immunologiczną i jej wagę można rozważać w dwóch kategoriach: osobniczej i populacyjnej. W pierwszym przypadku, dzięki omówionym powyżej zjawiskom, możliwe jest „uczenie się” przez układ immunologiczny człowieka odpowiednich reakcji na czynniki chorobotwórcze i osobnicze przeżycie. W drugim przypadku można założyć, że sprawnie funkcjonujący układ immunologiczny umożliwi przetrwanie całej populacji i giną osobnicy ze zbyt słabymi reakcjami obronnymi, natomiast przeważająca większość przeżywa.

2.3. Teoria sieci immunologicznej

Istotną kwestią dotyczącą odpowiedzi immunologicznej jest regulacja jej nasilenia tak, by podczas jej stymulacji nie dochodziło do nadmiernych, destrukcyjnych reakcji. Od niedawna podkreśla się rolę tak zwanych limfocytów regulatorowych (*Treg, regulatory T cells*), które działając bezpośrednio i pośrednio – m.in. za pomocą

czynników rozpuszczalnych (głównie cytokin) supresyjnych – tłumią reakcje immunologiczne. W regulacji odpowiedzi immunologicznej, przede wszystkim humoralnej, odgrywają również rolę tak zwane przeciwciała antyidiotypowe. Założenia tej regulacji zostały zaproponowane w połowie lat siedemdziesiątych zeszłego stulecia przez N. Jernego w postaci teorii sieci immunologicznej (*immune network theory*). N. Jerne uznał, że w częściach zmiennych przeciwciał i receptorów *BCR*, odpowiedzialnych za rozpoznanie antygeny, istnieją wzorce molekularne – tzw. idiotypy, których zespół tworzy unikatową dla każdego przeciwciała konformację (ukształtowanie) – idiotyp. Zaproponował, że konformacje te mogą indukować powstawanie antyprzeciwciał, określonych jako przeciwciała antyidiotypowe (*Ab2*) – rys. 3,



Rys. 3. Teoria sieci immunologicznej

Źródło: opracowanie własne.

regulujących (hamujących) stężenie tych przeciwciał, przeciwko którym są skierowane. Z kolei w stosunku do przeciwciał antyidiotypowych mogą się wykształcać przeciwciała antyantyidiotypowe (*Ab3*) – rys. 3, które w kolejnym kroku wygaszają odpowiedź immunologiczną itd. W ten sposób w sieci wzajemnych interakcji przeciwciał dochodzi do blokowania nadmiernej, niekontrolowanej odpowiedzi immunologicznej, na przykład przeciw antygenowi, który w sposób ciągły pobudza układ immunologiczny.

2.4. Teoria zagrożenia (immunologicznego)

Od wielu lat w dziedzinie immunologii zagadką było, dlaczego układ immunologiczny reaguje czasem opacznie na antygeny własne organizmu (autoantygeny), podczas gdy w niektórych sytuacjach, mimo obcości niektórych antygenów (np. pokarmowych), nie dochodzi do reakcji. Zjawiska te kłóciły się z powszechnie przyjętym założeniem, że warunkiem stymulacji odpowiedzi immunologicznej jest to, by antygen był obcy, natomiast w przypadku antygenów własnych dochodzi do tolerancji. Powyższe założenie wynikało z badań nad dojrzewaniem limfocytów *T* w grasicy. Podczas dojrzewania, w procesie negatywnej selekcji (*negative selection*), dochodzi

do eliminowania i „wyciszania” limfocytów *T*, które mogłyby rozpoznawać własne antygeny. Paradoks okazjonalnych reakcji na autoantygeny i tolerowania struktur obcych próbuje godzić sformułowana w połowie lat dziewięćdziesiątych zeszłego wieku teoria zagrożenia (*danger theory*), sformułowana przez badaczkę amerykańską P. Matzinger. Teoria ta postuluje, że układ odpornościowy „decyduje”, czy rozpoznać odpowiedź immunologiczną na dany antygen, czy nie reagować, w zależności nie od tego, czy jest to antygen własny (*self*), czy obcy (*non-self*), ale od tego, czy jest on postrzegany jako zagrażający, czy też jako nieszkodliwy. W ujęciu komórkowym wygląda to następująco. Limfocyty *T* (uczestniczące – jak opisano wcześniej – w odpowiedzi typu komórkowego i koordynujące odpowiedź humoralną) rozpoznają antygeny odpowiednio przetworzone przez komórki prezentujące antygen (*APC*). Samo rozpoznanie antygeny nie prowadzi jednak do wzbudzenia komórki; raczej wykształca się stan tolerancji. Aktywacja „dziewiczego” limfocytu zachodzi natomiast wówczas, gdy dostarczony zostanie jej dodatkowy sygnał (*second signal*) ze strony komórki *APC*. W zjawisku tym uczestniczą różne cząsteczki, przede wszystkim cząsteczka *CD28* na powierzchni limfocytu *T* i cząsteczka *CD80* (lub *CD86*) na komórce *APC*. Interakcja ta jest kontrolowana – zależna jest od liczby reagujących cząsteczek, a to z kolei uzależnione jest od stanu wzbudzenia komórki *APC*. Jeśli uzyska ona w mikrośrodkowisku (tkance) sygnał niebezpieczeństwa (*danger signal*), dojdzie do prezentacji antygeny i aktywacji limfocytu. Sygnał ten najczęściej dostarczany jest przez elementy patogennych mikroorganizmów (wirusów, bakterii itp.) w warunkach infekcji, ale także przez komórki własne organizmu, które doświadczyły „stresu” – dochodzi wówczas do zjawisk autoimmunizacyjnych, czyli nieprawidłowych (i niepotrzebnych) reakcji w stosunku do własnych antygenów. W rezultacie niszczone są własne tkanki i narządy, przykładowo w reumatoidalnym zapaleniu stawów (niszczona jest chrząstka) i stwardnieniu rozsianym (niszczona jest tkanka nerwowa w ośrodkowym układzie nerwowym). Na szczęście, dzięki silnym mechanizmom regulacyjnym, atakowanie własnych tkanek jest wyjątkiem, podczas gdy po kontakcie z patogenem – reakcje immunologiczne występują prawie zawsze.

3. Modele mechanizmów działania sztucznych układów immunologicznych wzorowane na mechanizmach działania biologicznego układu immunologicznego

Modele mechanizmów działania sztucznych układów immunologicznych należą do grupy modeli inteligencji obliczeniowej, podobnie jak sztuczne sieci neuronowe, algorytmy genetyczne i algorytmy ewolucyjne. Ich działanie jest inspirowane zasadami funkcjonowania układu immunologicznego wyższych kręgowców. Tworzone są modele matematyczne i komputerowe symulujące działanie rzeczywistych układów immunologicznych. Powstała w ten sposób dziedzina, która łączy immunologię, informatykę i technikę (*engineering*). Spośród pojęć używanych w immunologii do

najczęściej wykorzystywanych w dziedzinie tworzenia sztucznych mechanizmów inspirowanych jej dorobkiem należą pojęcia: „struktura własna” (*self*), „struktura obca” (*non-self*), „przeciwciało” (*antibody*), „antygen” (*antigen*). W przypadku sztucznych mechanizmów immunologicznych, tak jak w ich naturalnych pierwowzorach, komórka własna jest traktowana jako składnik prawidłowy, akceptowany przez organizm, przeciwciało – jako taki element, którego rolą jest wykrywanie elementów uznawanych za niewłaściwe, a antygen – jako element wykrywany i uruchamiający zadziałanie elementów odpowiadających przeciwciałom biologicznego układu immunologicznego.

3.1. Algorytmy mechanizmów działania sztucznych układów immunologicznych

Układ immunologiczny żywych organizmów można traktować w uproszeniu jako klasyfikator binarny, którego głównym zadaniem jest odróżnianie elementów „własnych” i „obcych”. Przyjmując, że element „obcy” symbolizuje rozwiązywany problem, można stworzyć ścisły opis tego problemu i skonstruować algorytm jego rozwiązania (por. [Wierchoń 2001, s. 3-4]).

Badania dotyczące mechanizmów naśladujących mechanizmy działania rzeczywistego układu immunologicznego można podzielić na teoretyczne i dotyczące zastosowań. Badania teoretyczne odnoszą się do opisów mechanizmów działania układu immunologicznego za pomocą modelowania matematycznego, dokonywania analiz zbieżności i złożoności algorytmów budowanych dla szczegółowego przedstawiania funkcjonowania tych mechanizmów. Te dotyczące zastosowań skupiają się na tworzeniu na podstawie algorytmów, oprogramowania komputerowego, przeznaczonego do wspomaganie w rozwiązywaniu konkretnych problemów. W ostatnich latach próbowano zastosować algorytmy odzwierciedlające mechanizmy działania biologicznego układu immunologicznego do rozwiązywania złożonych problemów obliczeniowych. Do głównych zastosowań należą: bezpieczeństwo komputera, filtry antyspamowe, optymalizacja wielokryterialna, klasyfikacja danych, uczenie maszynowe, wykrywanie anomalii. Rozwiązywanie problemów zapewniania bezpieczeństwa systemów, np. systemów informatycznych, jest chyba najbardziej wyraźnym, wręcz narzucającym się zadaniem, w którym przydatne mogą być sztuczne układy immunologiczne, które zawierałyby mechanizmy naśladujące funkcjonowanie układu immunologicznego żywych organizmów, aby podobnie jak one bronić w razie wrogiej inwazji czy też chronić przed czynnikami destrukcyjnymi. Tego typu algorytmy, inspirowane działaniem biologicznego układu immunologicznego, są próbą symulacji rozwiązań biologicznych i dokładnego odwzorowania funkcji polegającej na odróżnianiu elementów „własnych” od „obcych” i uruchomieniu ciągu działań, mających na celu usunięcie zagrożenia, co może być bardzo przydatne do zbudowania skutecznego komputerowego systemu zabezpieczeń. Inne zastosowania algorytmów naśladujących zdolność rozpoznawania przez układ immunologiczny żywych organizmów elementów „włas-

nych” i „obcych” to wykrywanie chorób (detekcja: chory, zdrowy), wad konstrukcyjnych (występowania elementów wadliwych lub ich braku), np. w konstrukcji mostów, stadionów sportowych, wirusów komputerowych (wykrywania, czy program jest wirusem, czy nim nie jest). Przykładem systemu antywirusowego, którego rozwiązania zostały zainspirowane działaniem biologicznego układu immunologicznego, może być *Digital Immune System* Symanteca [Symantec...].

Przykładami algorytmów realizujących funkcje bardziej odbiegające od funkcji, jaką pełnią biologiczne układy immunologiczne, a wzorowanych na działaniu tych układów, są algorytmy dotyczące zagadnień optymalizacji, w których element algorytmu odpowiadający patogenowi reprezentuje optimum globalne, a poszukiwany jest najlepiej dopasowany element reprezentujący przeciwcało. W problemach optymalizacji poszukiwane optimum może być bezpośrednio przyrównane do antygeny, który musi być rozpoznany przez układ immunologiczny, a znalezione rozwiązanie – do przeciwcała wskazującego antygen – w rozważanym przypadku optimum. Tego typu algorytmy są odwołaniem do rozwiązań biologicznego układu immunologicznego, mającym charakter metafory w rozumieniu poszukiwania w rozwiązaniach biologicznych inspiracji do tworzenia ściśle sformułowanych skończonej liczby kroków postępowania, których funkcje czy też przeznaczenie są dalekie od biologicznego pierwowzoru. Przykładowo, poza wymienionym już zagadnieniem rozwiązywania problemów optymalizacji, mechanizmy działania rzeczywistych układów immunologicznych umożliwiające rozpoznawanie struktur „własnych” i „obcych” mogą być inspiracją do tworzenia rozwiązań umożliwiających klasyfikację czy też analizę danych, kiedy postępowanie przebiega od potraktowania zbioru danych jako zbioru patogenów, poprzez ich rozpoznawanie i zapamiętywanie, aż do naśladowania działania sieci immunologicznej, co może zasugerować pomysł utworzenia grafu przedstawiającego relacje między danymi. W dziedzinie eksploracji danych (*Data Mining*) jako przykład można podać wyszukiwanie jednorodnych grup tematycznych w dużych zbiorach dokumentów tekstowych inspirowane mechanizmem tworzenia pamięci immunologicznej, zaś w zakresie systemów uczących się – rozwiązywanie zadań maszynowego uczenia się i wnioskowania w czasie rzeczywistym (por. wymienione w: [Płaskonka 2011]). Należałoby jeszcze wspomnieć o algorytmach, których celem budowy jest modelowanie działania rzeczywistego układu immunologicznego poprzez tworzenie modeli matematycznych czy też modeli komputerowych, aby zrozumieć funkcjonowanie rozwiązań biologicznych. Tego typu algorytmy mogą być budowane w celach badawczych i dydaktycznych.

Algorytm zaproponowany przez Y. Ishidę w roku 1990 w pracy [Ishida 1990] był jednym z pierwszych opublikowanych algorytmów naśladowujących mechanizmy działania biologicznego układu immunologicznego. Algorytm ten został w roku 1996 udoskonalony przez J.E. Hunta i D.E. Cooke’a [Hunt, Cooke 1996], tak że mógł znaleźć już szersze praktyczne zastosowanie. Ostatnie prace są skoncentrowane na budowie czterech rodzajów algorytmów odzwierciedlających działanie czterech różnych mechanizmów układu immunologicznego. Są to algorytmy:

(i) algorytm negatywnej selekcji, (ii) algorytm naśladujący działanie sieci immunologicznej, zgodnie z teorią zaproponowaną przez N. Jernego, uwzględniającą interakcje idiotypowe i antyidiotypowe (budowa tzw. sztucznej sieci immunologicznej), (iii) algorytm selekcji klonalnej, (iv) algorytm oparty na wykorzystaniu koncepcji teorii zagrożenia.

Algorytm negatywnej selekcji

Koncepcja algorytmu negatywnej selekcji została zainspirowana mechanizmem biologicznego układu immunologicznego, w którym dochodzi do usuwania tych komórek, które rozpoznają własne struktury. Najogólniej mówiąc, do algorytmu zostają wprowadzone jako dane wejściowe dwa zbiory: zbiór z wzorcami własnych struktur i zbiór złożony z losowo generowanych elementów reprezentujących przeciwciała. Na wyjściu otrzymuje się zbiór elementów – przeciwciał, nazywanych detektorami, ponieważ wykrywają czy też rozpoznają wyłącznie obce struktury. Już z powyższego narzuca się wniosek o możliwości zastosowania takich algorytmów do wykrywania nieprawidłowości (np. w systemach zapewniania bezpieczeństwa). Reasumując, należy stwierdzić, że działanie algorytmu polega na wygenerowaniu losowej populacji elementów – przeciwciał, wyznaczeniu stopnia dopasowania elementów ze zbioru elementów – przeciwciał do elementów ze zbioru wzorców własnych struktur, i usunięciu tych elementów – przeciwciał, których stopień dopasowania przekracza przyjęty stopień tolerancji, a jeżeli nie przekracza, dołączeniu do zbioru przeciwciał – detektorów obcych struktur. Algorytm negatywnej selekcji jest algorytmem wykorzystywanym do generowania elementów odpowiadających przeciwciałom biologicznego układu immunologicznego. Opierając się na takich algorytmach przedstawianych w literaturze, można opisać ich działanie w sposób następujący: początkowo zakłada się, że zbiór elementów odpowiadających przeciwciałom jest zbiorem pustym, a następnie w sposób iteracyjny w kolejnych krokach algorytmu te elementy są generowane, dopóki nie zostanie wytworzona przyjęta – uznana za zasadną w danym przypadku – ich liczba. Kolejne kroki algorytmu, służące do wygenerowania pożądanej liczby elementów odpowiadających przeciwciałom rzeczywistego układu immunologicznego, polegają na: wygenerowaniu w sposób losowy elementu, obliczeniu odległości między tym elementem a każdym z elementów odpowiadających pojęciu *self* rzeczywistego układu immunologicznego i, jeżeli odległość ta jest większa od założonego progu, dodaniu elementu odpowiadającego przeciwciału do zbioru elementów odpowiadających zbiorowi przeciwciał rzeczywistego układu immunologicznego. W uproszczeniu algorytm można zapisać w postaci podanego pseudokodu (w literaturze można znaleźć bardziej szczegółowe opisy algorytmu por. np. [De Castro 2006, s. 288-296; Dasgupta, Niño 2009, s. 62-110; Kołodziejczyk]):

Wejście:

S = zbiór wzorców własnych struktur.

Wyjście:

D = zbiór detektorów obcych struktur.

Powtarzaj aż do spełnienia kryterium stopu (pożądana liczba elementów zbioru D).

1. Wygeneruj losowo potencjalne detektory i umieść je w zbiorze P.

2. Określ podobieństwo (odległość) każdego elementu zbioru P i każdego elementu zbioru S.

3. Jeżeli dany potencjalny detektor $p \in P$ rozpoznaje przynajmniej jeden element $s \in S$ (ze względu na przyjęty próg podobieństwa), usuń detektor p z rozważań. W przeciwnym razie dodaj element p do zbioru D.

Algorytm negatywnej selekcji został zaproponowany w pracy [Forrest i in. 1994], w której przedstawiono możliwość komputerowej implementacji biologicznego mechanizmu negatywnej selekcji. Naśladując biologiczny pierwowzór, algorytm generuje zbiór detektorów rozpoznających określony wzorzec poprzez wykrycie faktu, że są one względem niego komplementarne. Algorytm i jego dalsze udoskonalone wersje najczęściej są stosowane do wspomagania rozwiązywania problemów detekcji anomalii i wykrywania wirusów komputerowych. Zastosowaniem algorytmów negatywnej selekcji jest także przeprowadzanie klasyfikacji.

Sztuczne sieci immunologiczne

W trakcie odpowiedzi immunologicznej zachodzi zjawisko jej wygaszania, zgodnie z teorią sieci immunologicznych N. Jernego (o której już pisano). Teoria ta stała się inspiracją koncepcji sztucznych sieci immunologicznych, które ujmują zjawiska interakcji przeciwciało–antygen (zwane interakcjami idiotypowymi) i przeciwciało–przeciwciało (zwane interakcjami antyidiotypowymi). Dla sztucznych sieci immunologicznych sugestią biologiczną jest zachodzące w biologicznym układzie zjawisko polegające na tym, że przeciwciała mogą pasować do innych przeciwciał tak dobrze jak do antygenów. Najogólniej algorytm działania takiej sztucznej sieci immunologicznej można przedstawić w następujący sposób [De Castro 2006, s. 303-312]. W sztucznej sieci immunologicznej naśladowany jest mechanizm zachodzenia interakcji idiotypowych i antyidiotypowych opierający się na wzajemnym oddziaływaniu limfocytów B . Działanie algorytmu rozpoczyna się od utworzenia zbioru elementów, odpowiadających populacji antygenów Ag (*antigens*). Zbiór ten stanowi zbiór uczący (zbiór wzorców, które będą rozpoznawane). Następnie generowany jest zbiór elementów odpowiadających populacji limfocytów B . Często limfocyty B są utożsamiane z przeciwciałami, co jest bardzo dużym uproszczeniem działania algorytmu w stosunku do biologicznego pierwowzoru. Elementy odpowiadające przeciwciałom są detektorami elementów – antygenów. Zbiór elementów reprezentujących w sztucznym systemie populację limfocytów B (przeciwciał) jest generowany w sposób losowy. W wielu algorytmach przeprowadza się podział zbioru uczącego (antygenów Ag) na dwa podzbiory, z których tylko jeden jest stosowany w procedurze uczenia elementów – limfocytów B , tak aby uzyskać odpowiednio działającą sztuczną sieć immunologiczną naśladującą działanie biologicznej sieci. Kolejnym krokiem algorytmu jest przeprowadzanie uczenia sztucznej sieci. Początkowo prze-

prowadzane jest uczenie, polegające na tworzeniu interakcji idiotypowych (tworzeniu sieci jedynie z możliwością zachodzenia interakcji idiotypowych). Dla każdego elementu należącego do zbioru odpowiadającego zbiorowi antygenów A_g wyznaczana jest wartość powinowactwa ze wszystkimi elementami ze zbioru elementów odpowiadających zbiorowi przeciwciał Ab (*antibodies*). Do określenia powinowactwa najczęściej wykorzystywana jest wybrana miara odległości (przeważnie euklidesowa). Zostaje utworzona macierz odległości zawierająca odległości pomiędzy elementami reprezentującymi przeciwciała a elementami reprezentującymi antygeny. Mniejsza odległość między elementem – przeciwciałem – a elementem – antygenem – jest odzwierciedleniem większego powinowactwa między przeciwciałem a antygenem biologicznej sieci. Następnie jest tworzony zbiór elementów reprezentujących przeciwciała o najmniejszej odległości (w sieci naturalnej – o największym powinowactwie) od każdego z antygenów. Realizowany jest proces klonowania (czyli tworzenia kopii) elementów ze zbioru elementów – przeciwciał. W procesie klonowania przyjmuje się założenie, że element – przeciwciało – o największym powinowactwie tworzy największą liczbę kopii. Przeciwciała będące klonami są poddawane mutacji, przy czym przeciwciała o największym powinowactwie podlegają najmniejszej mutacji. Po przeprowadzeniu mutacji powstaje nowy zbiór elementów – przeciwciał. Następnie wybierany jest ich podzbiór o największym powinowactwie z elementami – antygenami. Tworzą one podzbiór elementów odpowiadających komórkom pamięci. Następnie eliminowane są (analogicznie do układu biologicznego: śmierć komórek) te elementy – komórki pamięci ze zbioru, których wielkość powinowactwa przekracza założoną wartość. Dalej wyznaczane jest powinowactwo między klonowanymi przeciwciałami poprzez obliczenie odległości między reprezentującymi je elementami. Eliminowane są te elementy – klony ze zbioru komórek pamięci, dla których obliczone uprzednio powinowactwo jest poniżej zadanej wartości. Pozostałe elementy – przeciwciała – tworzą nowy zbiór komórek (przeciwciał), złożony ze zbioru przeciwciał – komórek pamięci i nowo utworzonych przeciwciał. Wybór elementów o „najmniejszym”, „największym” powinowactwie jest realizowany poprzez zbadanie, czy odległość między nimi jest większa, czy też mniejsza od przyjętego progu „usuwania” lub „przyjmowania” komórek do zbioru. W praktyce w algorytmach zwykle najpierw podejmowane są kroki prowadzące do utworzenia zbiorów elementów – zbiorów komórek pamięci (przeciwciał) – kolejno dla każdego z elementów – antygenów, tzn. dla każdego antygeny jest tworzony odrębny zbiór komórek pamięci (zbiór przeciwciał). Dopiero po utworzeniu takich zbiorów dla każdego antygeny należącego do zbioru wszystkich antygenów (populacji antygenów) w następnych krokach algorytmu najpierw wyznaczane jest powinowactwo między wszystkimi przeciwciałami pamięci, a następnie eliminowane są przeciwciała, między którymi powinowactwo leży poniżej pewnej wielkości nazywanej progiem supresji. Ostatecznie tworzony jest zbiór przeciwciał złożony ze zbioru przeciwciał – umownych komórek pamięci – i ze zbioru nowo utworzonych przeciwciał. W algorytmach sztucznych systemów immunologicznych elementy re-

prezentujące limfocyty B , które po przeprowadzeniu mutacji różnią się znacznie od siebie, są usuwane, co naśladuje działanie rzeczywistego układu immunologicznego. Poprzez takie usuwanie komórek biologiczny układ immunologiczny chroni się przed sytuacją powstawania komórek, które byłyby komórkami autoreaktywnymi. W algorytmach sztucznych sieci immunologicznych „znaczna różnica” między komórkami jest wskazywana poprzez wprowadzenie parametru określającego próg powinowactwa dla usuwania komórek, co oznacza, że po przekroczeniu przyjętej wartości tego progu takie elementy są eliminowane. W przypadku rzeczywistego układu immunologicznego żywych organizmów istnieje mechanizm powodujący hamowanie odpowiedzi immunologicznej. Ten mechanizm jest implementowany w algorytmach działania sztucznych sieci poprzez wprowadzenie dla każdego elementu odpowiadającemu antygenowi działania polegającego na usuwaniu tych elementów – komórek pamięci, między którymi zachodzi znaczne podobieństwo. W algorytmach do realizacji tego procesu wprowadza się wspomniany już parametr zwany progiem supresji. Jeżeli odległość między elementami odzwierciedlającymi komórki pamięci znajduje się poniżej założonej wartości tego progu, są one usuwane. Ostatecznym rezultatem działania przedstawionego algorytmu jest uzyskanie zbioru elementów, odpowiadającego zbiorowi przeciwciał Ab biologicznego układu, uzyskanych podczas działań, które można przyrównać do procesu uczenia sztucznej sieci. Kryteriami „stopu” dla przedstawionego rodzaju algorytmów może być wskazanie liczby iteracji czy też wskazanie dopuszczalnej wartości błędu podczas identyfikacji wzorców ze zbioru uczącego Ag . Zostały opracowane różne wersje przedstawionego algorytmu działania sztucznej sieci immunologicznej, także wersje z licznymi uproszczeniami (por. [Dasgupta, Niño 2009, s.120-147]). Przedstawmy jego działanie w postaci ogólnego pseudokodu [Wajs i in. 2005, s. 52-53; Kołodziejczyk]:

Pierwszym krokiem przedstawionego algorytmu jest załadowanie populacji antygenów. Jest to zbiór uczący, który zawiera wzorce. Drugim krokiem jest inicjalizacja populacji limfocytów B , która została tu utożsamiona z przeciwciałami mającymi odgrywać rolę receptorów dla antygenów.

Potem rozpoczyna się proces uczenia sieci immunologicznej. Dla każdego antygeny należące do zbioru uczącego przeprowadzamy następujące operacje:

- 1) wyznaczamy powinowactwo pomiędzy j -tym antygenem i wszystkimi przeciwciałami;
- 2) tworzymy zbiór A zawierający n przeciwciał o najwyższym powinowactwie;
- 3) n wybranych przeciwciał poddajemy procesowi klonowania proporcjonalnie do ich powinowactwa do antygenów; generujemy zbiór C klonów o najwyższym powinowactwie;
- 4) elementy zbioru C są poddawane mutacji i generowany jest zbiór C^* ; prawdopodobieństwo mutacji jest odwrotnie proporcjonalne do powinowactwa przeciwciał rodzicielskich do j -tego antygeny;
- 5) wyznaczamy powinowactwa pomiędzy antygenem j do każdego elementu zbioru C^* ;
- 6) ze zbioru C^* wybieramy $z\%$ przeciwciał, które mają największy wskaźnik powinowactwa, i tworzymy z nich komórki pamięci M_j ; „ z ” jest to parametr określający, jaki procent populacji przeciwciał będzie poddawany procesowi klonowania w procesie aktywacji antygeny;

- 7) eliminowanie (śmierć) wszystkich komórek pamięci M_j , dla których wartość powinowactwa $D_{k,j}$ spełnia zależność $D_{k,j} > s_d$ (s_d jest to próg usuwania komórek, parametr);
- 8) wyznaczenie powinowactwa między klonowanymi przeciwciałami (s_i, k);
- 9) supresja, tzn. eliminacja tych klonów ze zbioru M , dla których zachodzi zależność $s_i, k < s_s$ – próg supresji, parametr);
- 10) połączenie wszystkich przeciwciał, które są zapisane w zbiorze M_j ; łączymy zbiory $Ab\{m\}$ – zbiór przeciwciał, które są komórkami pamięci $Ab\{d\}$ – zbiór nowo utworzonych przeciwciał, $Ab\{m\} \cup Ab\{d\} = Ab$ – dostępny zbiór przeciwciał.

Algorytm selekcji klonalnej

Algorytmy selekcji klonalnej są przeznaczone do wybierania tych elementów sztucznych układów immunologicznych, które odpowiadają najbardziej skutecznym przeciwciałom biologicznych układów. Ogólna idea działania tych algorytmów polega na wybraniu elementów reprezentujących przeciwciała najlepiej dopasowanych do elementów reprezentujących antygeny i ich powieleniu (sklonowaniu). Zazwyczaj zakłada się, że powielanie jest proporcjonalne do stopnia dopasowania elementów – przeciwciał. Następnie przeprowadza się mutację elementów – przeciwciał, tak aby przynajmniej część z nich odznaczała się lepszym stopniem dopasowania niż przed mutacją. Na podstawie istniejących algorytmów selekcji klonalnej można przedstawić ogólny schemat działania takich algorytmów. W algorytmach zakłada się początkowo, że mamy pewien zbiór elementów odpowiadających przeciwciałom biologicznego układu immunologicznego i zbiór elementów odpowiadających antygenom (np. wygenerowanych w sposób losowy). W kolejnych krokach są realizowane następujące działania. Dla każdego elementu – antygeny – wybieramy założoną liczbę, najmniej odległych zgodnie z przyjętą metryką, elementów – przeciwciał. Następnie dla każdego wybranego elementu – przeciwciała – tworzymy liczbę klonów odwrotnie proporcjonalną do odległości (przyjmujemy założenie, że przeciwciała, które zidentyfikowało określony antygen, generuje największą liczbę kopii w porównaniu z przeciwciałami o mniejszej wartości powinowactwa). Następnie przeprowadzamy operację mutacji każdego klonu proporcjonalnie do odległości od rodzica (w ten sposób przeciwciała o największej wartości powinowactwa będzie miało najniższą wartość współczynnika mutacji, zostanie więc zmodyfikowane w mniejszym stopniu). Dołączamy klony do zbioru elementów odpowiadających przeciwciałom. Powtarzamy kolejne kroki algorytmu aż zostanie spełniony założony warunek „stopu”, np. brak elementu w zbiorze elementów odpowiadających antygenom, nierozpoznanego elementu – antygeny – lub osiągnięcie określonej liczby poprawnie rozpoznanych elementów – antygenów.

Podsumowując, można stwierdzić, że algorytm obejmuje następujące podstawowe kroki: (i) wybór zadanej liczby przeciwciał o największej wartości powinowactwa ze zbioru przeciwciał i utworzenie ich kopii proporcjonalnie do wartości powinowactwa do antygeny; przeciwciała o największej wartości powinowactwa posiada największą liczbę kopii; (ii) poddanie procesowi mutacji wszystkich nowo

utworzonych kopii przeciwciał ze wskaźnikiem mutacji odwrotnie proporcjonalnym do wartości powinowactwa; duża wartość powinowactwa przeciwciała do antygeny powoduje niską wartość wskaźnika mutacji; (iii) dodanie wszystkich zmodyfikowanych w poprzednim kroku przeciwciał do ich zbioru; (iv) kroki od pierwszego do trzeciego są powtarzane do momentu osiągnięcia założonego kryterium „stopu”. W postaci pseudokodu algorytm selekcji klonalnej można przedstawić następująco (por. szczegółowy opis w [De Castro 2006, s. 296-303; Dasgupta, Niño 2009, s. 117-120; Kołodziejczyk]):

Wejście:

- S – zbiór elementów – antygenów (wygenerowany w sposób losowy),
- n – liczba najgorszych elementów przeznaczonych do usunięcia.

Wyjście:

- M – zbiór detektorów zdolnych klasyfikować nieznanne antygeny.

1. Utworzyć w sposób losowy zbiór elementów – przeciwciał A.
2. Dla każdego antygeny ze zbioru S:
 - 2.1. Obliczyć odległość do każdego elementu – przeciwciała ze zbioru A.
 - 2.2. Utworzyć klony elementów – przeciwciał, które mają największe powinowactwo z antygenem.
 - 2.3. Liczba klonów jest odwrotnie proporcjonalna do odległości przeciwciała od antygeny.
 - 2.4. Przeprowadzić mutację klonów i umieścić elementy – przeciwciała o najwyższym powinowactwie do antygeny w zbiorze M.
 - 2.5. Usunąć ze zbioru A n elementów przeciwciał o najniższym powinowactwie do antygeny i wygenerować losowo n nowych elementów do zbioru A.

W roku 2000 w pracy [De Castro, von Zuben 2000] został zaproponowany algorytm selekcji klonalnej (*CSA – Clonal Selection Algorithm*), znany później pod nazwą *CLONALG (CLONal Selection ALGorithm)* [De Castro, von Zuben 2002]. Algorytm ten znalazł przede wszystkim zastosowanie w rozwiązywaniu problemów optymalizacji i rozpoznawania wzorców. Inspiracją dla opracowania algorytmu jest mechanizm selekcji klonalnej biologicznego układu immunologicznego, przy założeniu, że liczba klonów – przeciwciał, jest proporcjonalna do powinowactwa z antygenem, a rozmiar mutacji jest odwrotnie proporcjonalny, co jest zgodne z założeniami przyjętymi w przedstawionym, ogólnym schemacie algorytmów selekcji klonalnej. Zauważalne jest pewne podobieństwo algorytmów selekcji klonalnej do algorytmów genetycznych, lecz różnią się one od nich w realizacji wielu kroków algorytmu. Przykładowo w ogóle nie występuje operacja krzyżowania. W celu odwzorowania funkcjonowania mechanizmu selekcji klonalnej zostały opracowane także inne algorytmy niż wymieniony *CLONALG*. Algorytmy selekcji klonalnej znajdują zastosowanie przede wszystkim w optymalizacji wielomodalnej, rozwiązywaniu problemów kombinatorycznych (jak np. problem komiwojażera (*TSP – Traveling Salesman Problem*)).

Algorytm oparty na teorii zagrożeń

Pierwsze wersje algorytmu, ostatniego z wymienionych wcześniej rodzajów algorytmów, mianowicie algorytmu odzwierciedlającego koncepcję teorii zagrożenia (niebezpieczeństwa), powstały na podstawie propozycji zawartych w modelu przedstawionym przez P. Matzinger, opisanym w niniejszym artykule. Należą do mniej znanych i rzadziej stosowanych algorytmów niż dotychczas przedstawione: negatywnej selekcji, sieci immunologicznej, zgodnej z teorią sieci immunologicznej N. Jernego czy selekcji klonalnej. Z podstawami algorytmów opartych na koncepcji teorii zagrożenia można się zapoznać m.in. w: [Aickelin, Cayzer 2002; Dasgupta, Niño 2009, s. 153-160].

Łączenie algorytmów na potrzeby zastosowań

Algorytmy naśladujące działanie mechanizmów biologicznego układu immunologicznego są stosowane łącznie, np. kolejno: algorytm negatywnej selekcji, a następnie algorytm selekcji klonalnej. Przykładowo, jeżeli są używane algorytm negatywnej selekcji i algorytm selekcji klonalnej, to najpierw za pomocą algorytmu negatywnej selekcji tworzony jest zbiór elementów odpowiadających przeciwciałom rzeczywistego układu immunologicznego, a następnie wprowadzany jest element odpowiadający antygenowi rzeczywistego układu, który prowadzi do „infekcji” (element – antygen – może wprowadzić użytkownik w interakcji z programem realizującym algorytm). Następnie wykorzystywany jest algorytm selekcji klonalnej, który identyfikuje element – antygen „infekujący” – i „usuwa infekcję”. Użytkownik programu realizującego algorytmy ma zwykle możliwość sterowania jego działaniem za pomocą parametrów. W przypadku opisywanego stosowania algorytmu negatywnej selekcji i algorytmu selekcji klonalnej mogą być nimi przykładowo: liczba elementów pełniących rolę *self*, liczba generowanych elementów odgrywających rolę przeciwciał, liczba najbliższych elementów – przeciwciał – uaktywnianych przez element – antygen, liczba klonów tworzonych dla uaktywnionych elementów w roli przeciwciał, czas działania programu. Użytkownikowi mogą być prezentowane pośrednie, szczegółowe wyniki wykonywanego programu realizującego algorytm, jak np. elementy *self* błędnie rozpoznane przez elementy – przeciwciała. Różne elementy sztucznego układu immunologicznego, odpowiadające cząsteczkom układu immunologicznego żywych organizmów, jak np. przeciwciała czy antygen, bywają prezentowane użytkownikom przez programy komputerowe realizujące algorytmy za pomocą różnych kształtów i/lub kolorów.

Należy zaznaczyć, że również inne mechanizmy działania układu immunologicznego żywych organizmów omówione w części artykułu poświęconej temu układowi stały się inspiracją budowy algorytmów, będących podstawą dla twórców oprogramowania komputerowego. Przykładem może być model działania szpiku kostnego z jego zdolnością generowania populacji komórek i molekuł immunologicznych.

Zostały też opracowane tzw. algorytmy hybrydowe [Świtalska, rozdział 2]. Zalicza się do nich takie algorytmy, w których klasyczne algorytmy immunologiczne,

np. algorytm negatywnej selekcji i algorytm selekcji klonalnej, są łączone z innymi algorytmami sztucznej inteligencji, np. algorytmami genetycznymi. Algorytmy, zwane algorytmami immunogenetycznymi, wykorzystują operatory genetyczne: krzyżowania, mutacji, selekcji, przy zachowaniu odwzorowania podstawowych elementów układu immunologicznego, jak populacja antygenów i przeciwciał. Przykładowo algorytmy genetyczne używane są do generowania populacji przeciwciał, a następnie stosowana jest negatywna selekcja. Do określania powinowactwa w tych algorytmach wykorzystywane są często miary przyjmowane w algorytmach immunologicznych, opierające się na pomiarach stopnia dopasowania przeciwciała do antygeny. Innym przykładem algorytmu hybrydowego jest połączenie algorytmów realizujących mechanizmy inspirowane działaniem mechanizmów biologicznego układu immunologicznego z algorytmami sztucznych sieci neuronowych. Interesującym przykładem algorytmów zaliczanych do algorytmów hybrydowych są algorytmy immunologiczne z operatorem odzwierciedlającym zjawisko starzenia się komórek (*aging*). Rolą operatora jest naśladowanie mechanizmu śmierci komórek układu immunologicznego żywych organizmów. Pełni on funkcję uzupełniającą dla algorytmów selekcji komórek, które działają, opierając się na pomiarach stopnia dopasowania dla odzwierciedlenia biologicznego pojęcia powinowactwa. Operator jest stosowany w celu usuwania struktur, gdy uznaje się, że zbyt długo nie wykazują poprawy co do stopnia dopasowania, oraz w celu uniknięcia zbyt wczesnej zbieżności do optimum lokalnego. Oryginalną cechą algorytmu jest wprowadzenie pojęcia wieku komórki. Elementy przetwarzanych danych – komórki populacji, mają początkowo przypisany wiek równy 0. Gdy tworzone są kolejne generacje komórek, ich wiek jest zmieniany zgodnie z przyjętymi w danym algorytmie zasadami. Przyjmowane może być np. założenie, że komórki powstające jako klony mają wiek taki jak komórka, której są klonem, lub założenie, że jeżeli komórka zwiększa swój stopień dopasowania, to stale może być jej przyporządkowywany wiek równy 0. Komórki, które nie zwiększają stopnia dopasowania, „są coraz starsze”, ich wiek osiąga coraz wyższe wartości. Operator uwzględniający starzenie się komórek powoduje eliminację komórek, działając zgodnie z zasadą, że pora jej usunięcia jest zależna od wieku komórki, np. usuwa się komórki po przekroczeniu przez ich wiek założonej wartości progowej. Wprowadzane są ograniczenia, np. niedopuszczające do usuwania komórek odznaczających się wysokim stopniem dopasowania. Takie operatory, które nie dopuszczają do usuwania komórek z najwyższymi stopniami dopasowania, nazywane są elitarnymi (*elitist aging*). Operatory niezależni od procesu usuwania komórek od ich stopnia dopasowania nazywane są zwykłymi *pure aging*).

3.2. Reprezentowanie elementów składowych i mechanizmów sztucznego układu immunologicznego

Komórki i cząsteczki biologicznego układu immunologicznego charakteryzuje różnorodność ukształtowania przestrzennego. Aby zaprojektować sztuczny system mo-

delowania, przetwarzania i analiz danych, uwzględniający tę różnorodność, stworzono koncepcję, aby wykorzystać podejście polegające na odwzorowaniu komórek i cząsteczek za pomocą różnych struktur danych. Aby z kolei odwzorować zjawisko odpowiedzi immunologicznej – „wykrywania powinowactwa” (np. wykrycia antygeny przez przeciwciała), logiczny wydaje się pomysł badania stopnia komplementarności kształtów, tj. wzajemnego uzupełniania (jak np. wklęsły – wypukły). Inną możliwością może być rozważanie, zamiast komplementarności, stopnia podobieństwa w celu rozpoznania powinowactwa. Gdy rozpoznawanie antygeny jest mierzone za pomocą komplementarności, wówczas im większa „odległość”, tym większa zgodność co do dopasowania. Jeżeli rozpoznawanie jest mierzone za pomocą badania podobieństwa, to im mniejsza „odległość”, tym większa zgodność co do dopasowania.

W najwcześniejszych tworzonych algorytmach przeciwciała i antygeny były reprezentowane za pomocą łańcuchów (ciągów) binarnych. Biologiczne „relacje” przeciwciała i antygenów były odzwierciedlane za pomocą analizy częściowego dopasowania ciągów binarnych. Do tego celu często znajdowała zastosowanie odległość Hamminga (zob. rys. 4) i pomiar podobieństwa przy założonej wielkości tzw. parametru r (im mniejsza odległość, tym większe dopasowanie i powinowactwo) (por. np. [De Castro 2006, s. 282-284]).

1	1	1	1	1	0	0	1	0	0	0	1
1	0	1	1	0	1	0	1	1	1	1	0

Rys. 4. Odległość Hamminga, $r = 5$: ciągi binarne są jednakowe na pięciu pozycjach

Źródło: [Praczyk 2011].

Jeżeli zastosujemy miarę odległości Hamminga, to antygen (element „obcy”) zostanie zidentyfikowany, gdy odległość między ciągiem reprezentującym element „obcy” a ciągiem reprezentującym przeciwciało (lub tzw. detektor) jest mniejsza lub równa wartości parametru r . Stosowana bywa tzw. reguła *rcb* (*r-contiguous bits*) [De Castro 2006, s. 283], która zakłada, że dwa ciągi binarne są zgodne, gdy w każdym istnieje taki sam podciąg co najmniej długości r (por. rys. 5). Wtedy element obcy zostanie zidentyfikowany, jeżeli w zbiorze elementów – przeciwciała, istnieje co najmniej jeden element, którego ciąg jest taki sam na r pozycjach jak w ciągu reprezentującym element obcy (antygen).

Wynik porównania ciągów binarnych jest zależny od wartości parametru r . Jeżeli r jest równy długości porównywanych ciągów, każdy detektor (w rzeczywistym układzie immunologicznym – przeciwciało) może wykryć najwyżej jeden element

1	1	1	1	1	0	0	1	0	0	0	1
1	0	1	1	0	0	0	1	0	0	1	1

Rys. 5. Wykorzystanie reguły *rcb*, gdy przyjęto $r = 5$: ciągi binarne są takie same na pięciu kolejnych pozycjach

Źródło: [Praczyk 2011].

obcy. Jeżeli przyjmiemy, że r jest równe zero, to będziemy musieli uznać, że dowolne dwa ciągi binarne są zgodne, co oznacza w każdym porównaniu wykrycie obcego elementu. Im zatem większe wartości parametru r , tym mniejsza zdolność detektorów do wykrywania elementów obcych. Im mniejsze wartości r , tym większa liczba elementów obcych, które są możliwe do znalezienia za pomocą detektora. Przyjmując określoną wartość parametru r , musimy pamiętać, że ma ona wpływ na błędy w wykrywaniu elementów obcych. Jeżeli obiekt własny zostanie uznany za obcy, to wystąpi błąd nazywany *false positive*. Natomiast odwrotnie, jeżeli obiekt obcy zostanie rozpoznany jako własny, wówczas błąd nosi nazwę *false negative*. Oba rodzaje błędów mogą przynieść poważne niekorzystne skutki. W przypadku błędu określanego jako *false positive* (możemy stwierdzić, że mamy do czynienia z wystąpieniem „falszywego alarmu”) zostanie wyeliminowany element, który nie stwarza zagrożenia, ale może być nawet elementem mającym duże znaczenie dla usuwania elementów obcych. W przypadku błędu określanego jako *false negative* może dojść do niewykrycia elementu stwarzającego zagrożenie [Praczyk 2011].

Powinowactwo (*affinity*) między antygenem a przeciwciałem może być przedstawiane różnie, w zależności od tego, czy dopasowanie (*match*) między łańcuchem binarnym reprezentującym antygen a łańcuchem binarnym reprezentującym przeciwciało jest wyrażane za pomocą komplementarności (*complementarity*), czy za pomocą podobieństwa (*similarity*) (por. np. [De Castro 2006, s. 282-284]).

Rozważmy dwa przykładowe łańcuchy: $Ag = [111110010001]$ i $Ab = [101100010011]$, gdzie Ag oznacza antygen, a Ab oznacza przeciwciało.

Aby zbadać powinowactwo między tymi łańcuchami reprezentującymi antygen i przeciwciało, przeprowadźmy ich porównanie, jak to zilustrowano w dalszej części tekstu. Dopasowanie mierzone za pomocą komplementarności między dwoma łańcuchami oznaczmy jako $D(Ag, Ab)$, a dopasowanie mierzone za pomocą podobieństwa możemy wyrazić jako $L - D(Ag, Ab)$, gdzie L jest długością łańcuchów reprezentujących antygen i przeciwciało.

$$Ag = [111110010001],$$

$$Ab = [101100010011],$$

skąd dopasowanie pozycji w obu ciągach:

[101101111101],

gdy komplementarność: $D(Ag, Ab)$, to powinowactwo wynosi 3,
gdy podobieństwo: $L - D(Ag, Ab)$, to powinowactwo wynosi 9.

Mogą być stosowane inne, bardziej złożone miary dopasowania i powinowactwa.

Ciekawym przypadkiem jest sytuacja, gdy badamy powinowactwo dotyczące przykładowo jednego przeciwciała, ale dwóch antygenów [De Castro 2006, s. 284]. Możemy wówczas przyjąć założenie, że zachodzi ono tylko wtedy, gdy jest większe lub równe przyjętemu przez nas progowi, przyjmijmy ε , i które możemy określić jako próg powinowactwa. Wprowadzenie progu powinowactwa pozwala nam odzwierciedlić biologiczny mechanizm zjawiska reaktywności krzyżowej (*cross – reactivity*). Aby zilustrować rolę progu powinowactwa i możliwości odzwierciedlenia za jego pomocą zjawiska biologicznego reaktywności krzyżowej, gdy jak uprzednio D jest dopasowaniem mierzonym za pomocą komplementarności i jest wykorzystywane do określenia powinowactwa między łańcuchami bitowymi proporcjonalnie do ich stopnia podobieństwa $L - D$ (L jak uprzednio jest długością łańcuchów bitowych), posłużmy się przykładem:

$$Ab = [1001100010011],$$

$$Ag1 = [111110010001],$$

$$Ag2 = [011101100011],$$

Dopasowanie ($Ab, Ag1$): [101101111101]

Dopasowanie ($Ab, Ag2$): [001110001111].

Powinowactwo $Ab - Ag1$: 9

Powinowactwo $Ab - Ag2$: 7

Jeżeli próg reaktywności krzyżowej $\varepsilon = 9$, to antygen $Ag1$ został rozpoznany przez Ab , natomiast antygen $Ag2$ nie został rozpoznany. Powinowactwo w sztucznym układzie – między Ab i $Ag1$ – wynosi 9, a między Ab i $Ag2$ wynosi 7. Ponieważ przyjęliśmy próg dopasowania ε równy 9, przeciwciało Ab rozpoznało $Ag1$, lecz nie rozpoznało $Ag2$. Możliwe staje się odzwierciedlenie biologicznego zjawiska reaktywności krzyżowej – niektóre elementy – antygeny, mogą już być rozpoznane, podczas gdy inne jeszcze nie.

W opracowywanych algorytmach rozważa się często trzy zbiory ciągów binarnych. Pierwszy zbiór składa się z „dojrzałych” detektorów naśladowujących zachowanie przeciwciał biologicznego układu immunologicznego. Ich zadaniem jest rozpoznawanie wszystkich elementów, które się różnią od „własnych”. Wykrywanie elementów „obcych” jest realizowane przy wykorzystywaniu różnych algorytmów, np. takich jak opisane powyżej. Drugi zbiór ciągów binarnych składa się z tzw. detektorów „niedojrzałych”. Są to takie detektory elementów „obcych”, które są generowane w sposób losowy. Są one poddawane testowi, który bada, czy nie wskazują one elementów „własnych” jako „obcych”. Te spośród nich, dla których test wykazuje,

że wskazują elementy „własne” jako „obce”, są usuwane ze zbioru detektorów. Jeżeli test wypadnie pomyślnie, detektor może stać się detektorem „dojrzałym”. Trzeci zbiór to zbiór złożony z ciągów binarnych reprezentujących obiekty „własne”. Są one stosowane do testowania detektorów „niedojrzałych”. Powstawanie i usuwanie detektorów, w tym także „dojrzałych”, odbywa się nie jednorazowo, lecz w procesie o charakterze ciągłym. Dzięki temu algorytm detekcji elementów „obcych” (antygenów) przystosowuje swoje działanie do zmian, jakim podlegają elementy „własne” oraz „obce”.

Obecnie przeciwciała i antygeny są częściej niż za pomocą ciągów binarnych opisywane za pomocą wektorów liczb rzeczywistych, stanowiących punkty w przestrzeni o wymiarach wektorów. Zakłada się, że podobieństwo między antygenem a przeciwciałem jest wyrażane za pomocą odległości między dwoma punktami w przestrzeni [De Castro 2006, s. 282]. Do wyznaczenia stopnia dopasowania (powinowactwa) może służyć odległość euklidesowa:

$$D = \sqrt{\sum_{i=1}^{i=N} (Ab_i - Ag_i)^2},$$

gdzie Ab_i jest i -tą składową wektora $Ab = \langle Ab_1, Ab_2, \dots, Ab_N \rangle$ reprezentującego przeciwciało, a Ag_i – i -tą składową wektora $Ag = \langle Ag_1, Ag_2, \dots, Ag_N \rangle$ reprezentującego antygen. Ab i Ag są wektorami N -wymiarowymi. Stopień dopasowania, czyli powinowactwa pomiędzy przeciwciałem a antygenem, można obliczyć za pomocą metryki Minkowskiego, przy założeniu różnej wartości parametru p :

$$D = \left(\sum_{i=1}^{i=N} |Ab_i - Ag_i|^p \right)^{\frac{1}{p}}, \quad p \geq 1.$$

Dla $p = 1$, D nazywamy odległością Manhattan albo miejską:

$$D = \sum_{i=1}^{i=N} |(Ab_i - Ag_i)|.$$

Możemy wyznaczyć odległość, przyjmując, że jest nią iloczyn skalarny nieujemnych liczb rzeczywistych: $D = Ab \cdot Ag^T = \sum_{i=1}^{i=N} (Ab_i \cdot Ag_i)$.

Wydaje się, że wygodnym rozwiązaniem dla naśladownictwa mechanizmów biologicznego układu immunologicznego, które może dać korzyści, przede wszystkim ze względu na przejrzystość rozwiązania, jest wykorzystanie chyba najprostsze- go typu struktur danych, jakim jest łańcuch, ciąg znaków, jak to już zilustrowano na podanych przykładach. Jako łańcuchy mogą zostać przyjęte ciągi wartości rzeczywistych, całkowitych, binarnych czy też innych symboli.

W przedstawianiu elementów składowych sztucznego układu immunologicznego często przyjmuje się wiele uproszczeń, np. odwzorowując limfocyty B , zakłada się, że zawierają jedynie jedno przeciwciało ($B \text{ cell} = Ab$), a następnie operacje są realizowane na populacji przeciwciał. Element przedstawiający Ab w sztucznym

układzie immunologicznym jest traktowany jako zbiór cech stanowiących dziedzinę problemu. W przypadku zagadnień, takich jak optymalizacja, problemy kombinatoryczne, interpolacja funkcji, systemy zabezpieczeń, najczęściej można spotkać ich reprezentację w postaci wektora liczb rzeczywistych lub całkowitych, gdy rozpatrywane są obiekty systemów rozpoznawania wzorców – najczęściej wektorów binarnych, a w przypadku analiz eksploracji danych *Data Mining* – zazwyczaj list obiektów.

4. Przykłady zastosowań mechanizmów układu immunologicznego żywych organizmów w systemach przetwarzania komputerowego danych i rozwiązywania problemów

Aktualnie pojawia się coraz więcej prac (artykułów, rozdziałów w książkach naukowych, referatów konferencyjnych) donoszących o skutecznych, udanych zastosowaniach procedur, algorytmów, programów, których działanie opiera się na naśladowaniu mechanizmów układu immunologicznego żywych organizmów i jest inspirowanych funkcjonowaniem biologicznego układu immunologicznego (por. np. [Dasgupta, Niño 2009; De Castro 2006; Brabazon, M. O'Neill 2006; Dasgupta, Yu, Nino 2011]). Szczegółowe przedstawienie poszczególnych zastosowań bazujących na rozwiązaniach biologicznego układu immunologicznego wymagałoby przygotowania odrębnego artykułu lub nawet monografii. Dlatego w tym tekście ograniczymy się do krótkiego wymienienia niektórych z nich. Nawet krótki przegląd wspomnianych tu przykładów zwraca jednakże uwagę na różnorodność dziedzin i problemów, w których mogą być wykorzystywane rozwiązania układu immunologicznego człowieka.

Jedno z zastosowań, w których sięgnięto do rozwiązań biologicznego układu immunologicznego, stanowi przedstawiony w artykule [Chang, Yeh 2012] system dotyczący oceny kredytobiorców bankowych (*credit scoring analysis*). Zaproponowano klasyfikator klientów wskazujący na ich zdolność w spłacie kredytów. Opracowany klasyfikator oparto na sztucznych sieciach immunologicznych *AINE-based classifier* (*Artificial Immune Network – based classifier*). Do testowania i oceny algorytmu zostały użyte dwie bazy danych kredytobiorców pochodzące z repozytorium Uniwersytetu Kalifornijskiego w Irvine: australijska (zawierająca 690 rekordów, w tym 55% wniosków zaopiniowanych negatywnie) i niemiecka (zawierająca 1000 rekordów, w tym 30% wniosków odrzuconych). Celem było zaliczenie wniosku kredytowego do jednej z dwóch kategorii: zaakceptowany lub odrzucony. Klasyfikator „immunologiczny” identyfikował własne komórki (dobre kredyty) i obce komórki (złe kredyty). Porównując rzeczywistą klasyfikację wniosków kredytowych i otrzymaną przez algorytm immunologiczny, uzyskano stopę poprawnej klasyfikacji: 88,70% dla danych australijskich i 79,50% dla danych niemieckich. Uzyskane wyniki świadczą o tym, że zastosowany algorytm może konkurować z innymi algorytmami immunologicznymi, jak *AIRS* (*Artificial Immune Recognition System*),

i przewyższa algorytmy *SAIS* (*Simple Artificial Immune System*) i model Timmisa – sztuczny układ immunologiczny stosowany jako narzędzie skupiania (z uczeniem „bez nauczyciela” (*unsupervised*)) [Timmis, Neal, Hunt 2000].

Z ciekawym zastosowaniem mechanizmów biologicznego układu immunologicznego w problemach optymalizacji kombinatorycznej można zapoznać się w pracy [Świtalska]. W przypadku optymalizacyjnych systemów immunologicznych funkcję antygenu spełnia funkcja celu, będąca jednocześnie miarą dopasowania przeciwciała do rozwiązywanego problemu. W algorytmach najważniejszą rolę odgrywają mechanizmy mutacji i selekcji. Mutacja ma na celu wprowadzenie modyfikacji w przeciwciele, aby udoskonalić rozwiązania, które z założenia mają dążyć do optimum. Selekcja ma na celu dobór tych przeciwciał, które znajdują się najbliżej optimum. Istotnym problemem są optima lokalne, ponieważ proces może zbiegać do optimum lokalnego, nie znajdując rozwiązania globalnego. Zapobiega się temu przez utrzymanie odpowiedniego stopnia zróżnicowania przeciwciał. Jest to realizowane przez mechanizmy selekcji, które zapobiegają dominacji populacji przez komórki podobne. Jednym z problemów optymalizacji kombinatorycznej jest problem kolorowania grafu. Problem kolorowania grafu $G = (V, E)$, gdzie V – zbiór wierzchołków, E – zbiór krawędzi, polega na przypisaniu K kolorów do wierzchołków grafu, żeby żadne dwa sąsiadujące (tj. połączone krawędzią) wierzchołki nie miały tego samego koloru. K jest pewną liczbą mniejszą od liczby wierzchołków lub jej równą. Zadanie polega więc na znalezieniu funkcji $f: V \rightarrow \{1, \dots, K\}$, żeby $f(v_1) \neq f(v_2)$ dla wierzchołków sąsiednich v_1, v_2 . Minimalna liczba kolorów $\chi(G)$, dla której istnieje rozwiązanie, nazywa się liczbą chromatyczną grafu G . Do rozwiązania problemu zastosowano hybrydowy algorytm immunologiczny z operatorem uwzględniającym starzenie się komórek. W przeprowadzonych eksperymentach warunkiem była znana wartość liczby chromatycznej grafu lub limit liczby generacji. Szybkość rozwiązywania problemu była szacowana na podstawie liczby generacji potrzebnych do osiągnięcia znanej wartości liczby chromatycznej. Efektywność była mierzona odsetkiem przebiegów, w których została znaleziona wartość liczby kolorów w zadanym limicie generacji. Efektywność mieściła się w przedziale 90-100%. Badania przeprowadzono na różnych grafach. Najprostsze miały 10 wierzchołków, 20 krawędzi i liczbę chromatyczną równą 4. Najbardziej złożone miały 128 wierzchołków, 5198 krawędzi i liczbę chromatyczną równą 73.

Innym przykładem zastosowania dorobku immunologii w obszarze spoza tej dziedziny jest opisana w pracy [Wajs i in. 2005] klasyfikacja wielowymiarowych danych medycznych. Zadaniem było przeprowadzenie klasyfikacji parametrów krwi tętnicznej. Początkowo antygeny i przeciwciała były modelowane za pomocą łańcuchów binarnych. Później stosowano ich reprezentację za pomocą ciągów liczb rzeczywistych. Zastosowano algorytm selekcji klonalnej. Wykorzystano także sieci immunologiczne działające zgodnie z teorią sieci N . Jernego z fazą trenowania sieci i fazą jej testowania. Rozpatrzono problem dotyczący diagnozowania w neonatologii, czyli w dziedzinie medycyny dotyczącej noworodka. Przeprowadzono badania

dotyczące przebiegu niewydolności oddechowej za pomocą parametrów gazometrycznych krwi tętniczej. Wykorzystano dane pochodzące z komputerowej bazy danych – Neonatal Information System (NIS). Obejmuje ona dane dzieci z Oddziału Intensywnej Terapii Noworodka Polsko-Amerykańskiego Instytutu Pediatrii Collegium Medicum Uniwersytetu Jagiellońskiego. Dane zbierane były przez kilka lat. W pracy porównano wyniki działania algorytmu dla różnych parametrów (np. liczba przeciwciał poddanych procesowi klonowania, próg usuwania klonów ze zbioru komórek pamięciowych).

W pracy [Laurentys i in. 2010] opisano rozwiązania zainspirowane mechanizmami działania biologicznego układu immunologicznego, dotyczące wykrywania usterek (wad, defektów) (*fault detection*). Wykrywanie usterek jest ważnym problemem, szczególnie w dziedzinie procesów inżynierskich. Jest jednym ze składników zarządzania nieprawidłowymi zdarzeniami (*AEM – Abnormal Event Management*). *AEM* dotyczy wykrywania, diagnozy i korekty nieprawidłowych warunków procesów w czasie rzeczywistym, przewidywania konsekwencji błędów. Procesy przemysłowe są bardzo złożone. Podano przypadki w literaturze, w których obserwuje się do 1500 zmiennych w ciągu sekundy. Taka liczba zmiennych oddziałujących ze sobą dynamicznie, pomimo automatyzacji, wymaga ingerencji człowieka. Ludzie odpowiedzialni za *AEM* podejmują często niepoprawne decyzje. Statystyki przemysłowe pokazują, że 70% wypadków jest spowodowanych przez błędy ludzi. Jest konieczne stworzenie narzędzia, które ma pomóc w monitorowaniu odchyleń od procedur, rozwiązywaniu problemów bezpieczeństwa i jakości procesów. Przydatne okazują się tu techniki oparte na immunologii. W artykule [Laurentys i in. 2010] opisano zastosowanie *AIS* do automatycznego monitorowania i wykrywania usterek w celu uniknięcia nieprawidłowego działania procesu i uniknięcia wypadków. Opisano dwa istniejące algorytmy oparte na negatywnej selekcji i zaproponowano nowy algorytm nazwany *Multi-Operational (MO) Algorithm*. Zaimplementowane dwa algorytmy negatywnej selekcji w celu wykrywania anormalnych zdarzeń to: algorytm stałego promienia (*Constant Radius Algorithm*) oraz algorytm zmiennego promienia (*Variable Radius Algorithm*). Algorytm stałego promienia polega na wygenerowaniu zbioru własnych komórek (*self set S*), promienia detekcji (*detector radius r*) i liczby detektorów (*m*), a następnie utworzenia losowego zbioru detektorów (*D*). Tworzy się nowy detektor $x \in [0;1]^n$. Dla każdego $s \in S$ obliczamy odległość $\rho(s, x) = d$. Jeżeli $d < r$, to tworzymy nowy detektor (wracamy do punktu tworzenia nowego detektora), a w przeciwnym razie dołączamy element x do zbioru *D*. Tak postępujemy aż do momentu, gdy liczba elementów zbioru *D* wyniesie *m*. Jeżeli jest mniejsza od *m*, wracamy do etapu tworzenia nowego detektora. Ten algorytm nie jest zdolny wyznaczyć granicy między *self* i *nonself*. Algorytm zmiennego promienia jest rozwinięciem poprzedniego algorytmu. Różnica polega na tym, że promień detekcji zmienia się, a także na tym, że do zbioru detektorów są dodawane tylko te elementy, które są niewykrywane przez inne detektory. Ten algorytm pozwala wyznaczyć granicę między *self* i *nonself*, ale ma tę wadę, że nie gwarantuje pokrycia zbioru

obcych komórek (*nonsel*f). To skłoniło do poszukiwania lepszych algorytmów i jednym z nich jest zaproponowany *MO Algorithm*. Jego celem jest znalezienie lepszego pokrycia zbioru „obcych” komórek i zwiększenie efektywności. Rozumiemy przez to mniejszą liczbę detektorów i zwiększenie promienia. Jednocześnie detektory powinny w małym stopniu zachodzić na siebie (*overlap*). W pracy porównano algorytm *MO* z dwoma wspomnianymi uprzednio algorytmami stosowanymi do wykrywania usterek. Nowy algorytm okazał się bardziej efektywny, zauważono jednak, że bardzo ważnym czynnikiem jest poprawna normalizacja danych rzeczywistych. Niepoprawna normalizacja utrudnia wykrywanie usterek, ponieważ różnica między danymi z usterkami a bez usterek jest bardzo subtelna i algorytm może nie być zdolny to wykryć.

5. Zakończenie

Immunologia jest obecnie bardzo szybko i intensywnie rozwijającą się dziedziną wiedzy. Świadczy o tym lawinowy wzrost publikacji z tej dziedziny: książek, artykułów w czasopiśmie naukowych, referatów konferencyjnych. Choć jest dziedziną należącą do nauk medycznych i rozwijającą się w ich ramach, gdzie jej przydatność w zakresie badań teoretycznych i zastosowań klinicznych nie budzi wątpliwości, coraz częściej do jej dorobku sięgają praktycy spoza nauk medycznych. Przykładem są opisywane w niniejszym artykule problemy modelowania, przetwarzania i analiz danych. Jak już wspomniano we wstępie, stała się ona inspiracją dla rozwoju nowego podejścia do obliczeń komputerowych (*immunocomputing*), jak i budowy sztucznego układu immunologicznego (*artificial immune system*). W obu przypadkach, różniących się nieco podejściem do sposobu jej zastosowania, wykorzystywane są coraz szerzej mechanizmy immunologiczne, i to nie tylko przez proste naśladowanie, ale i coraz częstsze wzbogacanie o nowe elementy. O wzroście znaczenia i rozwoju immunologii dla zastosowań w rozwiązywaniu problemów spoza jej naturalnego obszaru, jakim jest medycyna, można się przekonać, śledząc publikacje cyklicznie odbywających się konferencji *ICARIS* (*International Conferences on Artificial Immune Systems*) i dotyczące tworzenia oraz wykorzystywania mechanizmów sztucznie konstruowanych na podstawie naturalnych, biologicznych odpowiedników.

Aby wykorzystać mechanizmy układu immunologicznego żywych organizmów kształtujące się w procesie ewolucji, stanowiące złożony i skomplikowany układ biologiczny i zachodzące dzięki przemianom biologicznym, chemicznym w obszarze modelowania, przetwarzania i analiz danych, konieczne jest zaangażowanie wiedzy matematycznej, informatycznej, inżynierskiej. Muszą być budowane algorytmy matematyczne (informatyczne) naśladowujące, a także rozszerzające działanie mechanizmów układu immunologicznego żywych organizmów. Algorytmy te są z powodzeniem stosowane do rozwiązywania różnorodnych problemów, w szczególności problemów sztucznej inteligencji i *Data Mining*.

Wiele mechanizmów funkcjonowania układu immunologicznego żywych organizmów jest już obecnie przedstawianych w postaci modeli, algorytmów, procedur. Liczne są opisywane w literaturze.

Działania biologicznego układu immunologicznego mogą stanowić inspirację i nadają się do mniej lub bardziej dokładnego odwzorowania w postaci algorytmów przetwarzania komputerowego danych i rozwiązywania problemów z różnych dziedzin. Ze względu na ograniczoną objętość niniejszego artykułu przedstawiono w nim jedynie reprezentatywne – takie, które odznaczają się szczególnie wyrazistą analogią do mechanizmów spotykanych w biologicznym układzie immunologicznym. Należą do nich algorytmy oparte na: (i) negatywnej selekcji służącej definiowaniu zbioru detektorów umożliwiających wykrywanie anomalii, (ii) działaniu sieci immunologicznych z ich zdolnością do „lawinowego” rozpoznawania przeciwciał i zwiększania skuteczności zwalczania antygenów, (iii) selekcji klonalnej sprawiającej, że aktywacji ulegają jedynie wybrane komórki, przeciwciała – te które rozpoznają określone antygeny, powielają się, zwiększają odporność na patogen, niszczą go, by potem ulec apoptozie, czyli naturalnemu procesowi zaprogramowanej śmierci już „zużytej” komórki, wytwarzając jednocześnie pamięć immunologiczną zwiększającą odporność, (iv) koncepcji teorii zagrożenia.

Podstawowe zalety sztucznych konstrukcji czy też procedur immunologicznych rozpatrywane z obu wymienionych na wstępie punktów widzenia, tj. zarówno przetwarzania komputerowego, jak i opracowywania sposobów rozwiązywania problemów, to: uniwersalność umożliwiająca rozwiązywanie problemów o różnym stopniu złożoności, stwarzanie możliwości przeszukiwania wielowymiarowej – złożonej przestrzeni rozwiązań, łatwość implementacji, umożliwienie tworzenia algorytmów równoległych (por. w pracy: [Płaskonka 2011], gdzie wymieniono podstawowe z nich, jak się wydaje, w sposób kompletny). Można przytoczyć także pewne wady zastosowań rozwiązań inspirowanych dorobkiem immunologii: (1) brak gwarancji, że zostanie osiągnięte optimum w określonym czasie, (2) konieczność zindywidualizowanego dostosowania algorytmu dla każdego rozpatrywanego problemu (przede wszystkim w zakresie parametrów, ponieważ jeżeli zostaną ustalone niepoprawnie działanie algorytmu może nie sprostać oczekiwaniom), (3) nieistnienie opracowanego aparatu matematycznego, który umożliwiłaby ścisłą analizę postępowań naśladujących mechanizmy układów odpornościowych żywych organizmów (por. też [Płaskonka 2011]).

Dotychczasowe doświadczenia, także opisane w tym artykule, wskazują, że dorobek immunologii nie powinien pozostawać poza zasięgiem uwagi badaczy i praktyków zajmujących się modelowaniem, przetwarzaniem i analizami danych, ale winien być przedmiotem ich bacznej uwagi ze względu na rolę, jaką może odegrać w ich rozwoju, nasuwając pomysły nowych sposobów działania (algorytmów) i nowych zastosowań. W celu uzupełnienia, rozwinięcia i poszerzenia informacji zawartych w naszym krótkim artykule, czytelników zainteresowanych historią powstania i rozwoju „sztucznych układów immunologicznych”, współczesnymi kierunkami

ich teoretycznego rozwoju, jak i zastosowaniami, a nie dysponujących aktualnie czasem nad odrębnymi studiami obszerniejszych pozycji zachęcamy do zapoznania się z artykułem przeglądowym [Dasgupta, Yu, Nino 2011].

Literatura

- Aickelin U., Cayzer S., *The Danger Theory and Its Application to Artificial Immune Systems*, Proceedings of the 1st Internal Conference on Artificial Immune Systems (ICARIS-2002), s. 141-148, Canterbury 2002.
- Brabazon A., O'Neill M., *Biologically Inspired Algorithms for Financial Modelling*, Springer-Verlag, Berlin, Heidelberg 2006.
- Chang S.-Y., Yeh T.-Y., *An artificial immune classifier for credit scoring analysis*, "Applied Soft Computing" 2012, 12, s. 611-618.
- Dasgupta D., Niño L.F., *Immunological Computation. Theory and Applications*, Taylor & Francis Group, LLC, (Auerbach/CRC is an imprint of Taylor & Francis Group), Boca Raton, London, New York 2009.
- Dasgupta D., Yu S., Nino F., *Recent advances in artificial immune systems: Models and applications*, "Applied Soft Computing" 2011, 11, s. 1574-1587.
- De Castro L.N., *Fundamentals of Natural Computing. Basic Concepts, Algorithms, and Applications*, Chapter 6: *Immunocomputing*, Taylor & Francis Group, LLC, (Chapman & Hall/CRC is an imprint of Taylor & Francis Group), Boca Raton, London, New York 2006.
- De Castro L.N., Von Zuben F.J., *Learning and optimization using the clonal selection principle*, "IEEE Transactions on Evolutionary Computation" 2002, 6 (3), s. 239-251.
- De Castro L.N., Von Zuben F.J., *The Clonal Selection Algorithm with Engineering Applications*, Genetic and Evolutionary Computation Conference GECCO'00 – Workshop Proceedings, Las Vegas 2000.
- Forrest S., Perelson A.S., Allen L., Cherukuri R., *Self-Nonself Discrimination in a Computer*, IEEE Symposium on Research in Security and Privacy, Los Alamos 1994.
- Gołąb J., Jakóbiński M., Lasek W., Stokłosa T. (red.), *Immunologia*, Wydawnictwo Naukowe PWN, wyd. 6, Warszawa 2012.
- Hunt J.E., Cooke D.E., *Learning using an artificial immune system*, "Journal of Network and Computer Applications" 1996, 19, s. 189-212.
- Ishida Y., *Fully Distributed Diagnosis by PDP Learning Algorithm: Towards Immune Network PDP Model*, IEEE International Joint Conference on Neural Networks, San Diego 1990.
- Kołodziejczyk J., *Obliczenia z wykorzystaniem sztucznej inteligencji*, wykład V – *Sztuczne systemy immunologiczne*, http://wikizmsi.zut.edu.pl/uploads/7/74/OzWSI_I_S1_W5.pdf [dostęp: 29.12.2012].
- Lasek W., *Immunologia. Podstawowe zagadnienia i aktualności*, wyd. 2, Wydawnictwo Naukowe PWN, Warszawa 2009.
- Laurentys C.A., Ronacher G., Palhares R.M., Caminhas W.M., *Design of an artificial immune system for fault detection: A negative selection approach*, "Expert Systems with Applications" 2010, 37, s. 5507-5513.
- Płaskonka J., *Zastosowanie sztucznych systemów immunologicznych w zagadnieniach optymalizacji*, 2011, http://www.ioz.pwr.wroc.pl/pracownicy/kuchta/Joanna_Plaskonka_Immune_Systems.pdf [dostęp: 28.12.2011].
- Praczyk T., *Zastosowanie sztucznych systemów immunologicznych do wykrywania ech radarowych o niskim poziomie sygnału*, „Logistyka” 2011, 6, s. 3461-3467.

- Symantec. *The Digital Immune System. Enterprise-Grade Anti Virus Automation in the 21st century, Technical Brief*, <http://www.symantec.com/avcenter/reference/dis.tech.brief.pdf> [dostęp: 28.12.2012].
- Świtalska A., *Sztuczne systemy immunologiczne – zastosowanie w optymalizacji kombinatorycznej*, <http://www.ipipan.waw.pl/~stw/ais/ks/> [dostęp: 28.12.2012].
- Timmis J., Neal M., Hunt J., *An artificial immune system for data analysis*, "BioSystems" 2000, 55, s. 143-150.
- Wajs W., Świącicki M., Wais P., Wojtowicz H., Janik P., *Zastosowanie sieci immunologicznej do drążenia danych. Data Mining with Artificial Immune System*, "Bio-Algorithms and Med-Systems. Journal edited by Medical College – Jagiellonian University" 2005, vol. 1, no. 1/2, s. 51-56.
- Wierzchoń S.T., *Sztuczne systemy immunologiczne. Teoria i zastosowania*, Akademicka Oficyna Wydawnicza EXIT, Warszawa 2001.

FROM IMMUNOLOGY TO MODELING, PROCESSING AND ANALYSIS OF DATA

Summary: Observation of nature is often the inspiration for unconventional and innovative ideas, proposals, suggestions, overall concepts for carrying out modeling phenomena and technical processes, economic, processing and analysis of statistical data. Artificial neural networks, genetic algorithms and evolutionary algorithms were created as an imitation of biological solutions. In recent years, the rapid development of immunological algorithms which are based on the action of the human immune system, has occurred. These algorithms are used in mathematics, engineering, problem solving in decision making, management, economics, finance. At the beginning of this article we describe the action of the human immune system, which is the basis for the construction of immune algorithms. Then we present the most important mechanisms influencing the operation and construction of the algorithms based on biological mechanisms of the immune system, in particular the negative selection, immune network, clonal selection. Then we describe how the elements and mechanisms of artificial immune system algorithms are represented to enable modeling, processing and analysis of data. In the last part of the article to illustrate the wide range of possibilities of using immune algorithms to solve various sorts of problems from a completely different fields, we present examples of applications that involve the assessment of bank borrowers, the task of coloring a graph belonging to combinatorial optimization problems, a multi-dimensional classification, fault detection.

Keywords: immunology, immunocomputing, artificial immune system, algorithms of artificial immune systems, representation of the elements of biological systems in artificial immune systems, the use of artificial immune algorithms in practice.