PROFUNDERE SCIENTIAM

nr 15
styczeń 2020

BIULETYN CENTRUM STUDIÓW ZAAWANSOWANYCH POLITECHNIKI WARSZAWSKIEJ

Cyber(nie) bezpieczeństwo

Krzysztof Szczypiorski

Streszczenie

Celem artykułu jest przedstawienie trzech tez: kulturowej → cyberprzestrzeń jest zaawansowanym tworem techniczno-kulturowym – jest realizacją marzeń wielu twórców, wynalazców i inżynierów; technicznej → bezpieczeństwo i cyberprzestrzeń to nierozłączne elementy (stąd: cyberbezpieczeństwo) i paranoicznej → pełne bezpieczeństwo, jeśli jest osiągalne, nie jest stanem stałym (stąd: cyber(nie) bezpieczeństwo). Cyberprzestrzeń jest rozumiana jako zbiór technik cyfrowych służących do wymiany informacji, ale także jako nowego typu przestrzeń społeczna częściowo wirtualna, która może być bytem całkowicie odseparowanym od fizycznego. Geneza nazwy: w latach 1968-1970 duńska artystka Susanne Ussing we współpracy z duńskim architektem Karstenem Hoffem stworzyła serię kolaży pod tytułem „CYBERSPACE". Dekadę później termin pojawił się w literaturze (William Gibson). Tradycyjnie ludzie próbowali rozumieć świat przez relacje pomiędzy różnymi zjawiskami fizycznymi zachodzącymi w otoczeniu np. poprzez żywioły. Cyberprzestrzeń jest tworem sztucznym jednak posiada związki z otoczeniem fizycznym – można ją traktować jako kolejny żywioł. Jako cezurę powstania cyberprzestrzeni można podać rok 1968, w którym pojawił się *routing* w sieci ARPANET, a także pierwszy programowalny sterownik logiczny (PLC). Natomiast dla cyberbezpieczeństwa będzie to rok 1976 – opublikowanie algorytmu uzgadniania klucza przez Witfielda Diffiego oraz Martina Hellmana. Rozwój bezpieczeństwa jest skorelowany z działaniami wojennymi i zbrojeniem – branża wojskowa dokonywała historycznie największych inwestycji w tym obszarze. W dalszej części wykładu przedstawiono cyfryzację mowy oraz wybrane ataki (podsłuch, modyfikacja, podszycie się i wyparcie). Określono podstawowe związki pomiędzy podstawowymi usługami cyberbezpieczeństwa: poufnością, integralnością, uwierzytelnieniem i niezaprzeczalnością. Przedstawiono związki pomiędzy zagrożeniem, podatnością, zasobami i ryzykiem, a następnie zaprezentowano generycznie projektowanie zabezpieczeń jako iteracyjny proces zawierający analizę ryzyka, projekt polityki bezpieczeństwa i oszacowanie kosztów. W dalszej części przedstawiono zagadnienia rozpoznawania znanych ataków i anomalii tłumacząc złożoność tworzenia modeli zachowania w cyberprzestrzeni, które przeważnie nie odzwierciedlają wszystkich cech i ich dynamizmu. Zagadnienie fałszywych alarmów wpływa na rzetelność systemów wykrywających anomalie, w szczególności błędy drugiego rodzaju prowadzą do nierozpoznawania ataków.

1. Wstęp

W toku ewolucji ludzkość rozwinęła swoje możliwości poznawcze i przystosowawcze dążąc do ekspansji w

{CIĄG DALSZY NA S. 6}

W NUMERZE

między innymi:

KWANTECHIZM czyli klatka na ludzi (fragm.)

Andrzej Dragan

Obrót czasoprzestrzeni

HISTORIA ŻYCIA NA ZIEMI to jakieś parę miliardów lat. Historia naszej cywilizacji to najwyżej kilkadziesiąt tysięcy lat. Oznacza to, że wkład, jaki wnosimy w historię życia na naszej planecie (będącej skądinąd pyłkiem na skraju jednej z dziesięciu miliardów galaktyk), jest mniej więcej taki, jak wkład rozdeptanej muchy umieszczonej na czubku Pałacu Kultury do jego wysokości.

Z punktu widzenia Ziemi, którą matematyk Hugo Steinhaus nazywał „kulą u nogi" gatunek ludzki jest więc najwyżej niesfornym epizodem. Podczas

{CIĄG DALSZY NA S. 19}

DZIAŁALNOŚĆ CENTRUM STUDIÓW ZAAWANSOWANYCH PW

czyli najważniejsze wydarzenia i najbliższe plany CSZ PW

PROFESOROWIE WIZYTUJĄCY

W minionym roku akademickim gościliśmy 13 wybitnych uczonych z ośrodków naukowych na całym świecie:

- prof. Michaił Żytomirski, Technion–Israel Institute of Technology, Haifa, Izrael;
- prof. Don Zagier, Max Planck Institute for Mathematics in Bonn, Niemcy;
- prof. Jun-Muk Hwang, Korea Institute for Advanced Study (KIAS), Korea Południowa;
- prof. Takashi Nishimura, Institute of Environment and Information Sciences, Yokohama National University, Japonia;
- prof. Shuichi Izumiya, Hokkaido University w Sapporo, Japonia;
- prof. Takuo Fukuda, Department of Mathematics, Nihon University, Japonia;
- prof. JJ Garcia-Luna Aceves, Computer Science and Engineering at the University of California, Santa Cruz (UCSC), USA;
- prof. Rolf Jeltsch, Departament Matematyki na ETH Zurich (Eidgenössische Technische Hochschule Zürich), Szwajcaria;
- prof. Goo Ishikawa, Hokkaido University, Japonia.

Profesorowie wizytujący z Europejskiego Towarzystwa Matematycznego:

- prof. Volker Mehrmann, Uniwersytet Techniczny w Berlinie, Instytut Matematyki MA, Niemcy;
- prof. Armen Sergeev, Instytut Matematyczny Steklova (Rosyjska Akademia Nauk), Rosja;
- prof. Roland Duduchava, Ivane Javakhishvili State University, Tbilisi, Gruzja;
- prof. Alice Fialowski, University of Pécs, Eötvös Loránd University, Węgry.

KONWERSATORIUM PW

Odbyły się 3 odczyty specjalne w ramach Konwersatorium Politechniki Warszawskiej:

Semestr zimowy 2018/2019

29 listopada 2018 r. – Profesor Utz von Wagner z Uniwersytetu Technicznego w Berlinie, odczyt pt. *Nonlinear Oscillations in Mechanical and Mechatronic Systems: Set-ups, Methods, Phenomena and Technical Examples*.

Semestr letni 2018/2019

23 maja 2019 r. – Profesor Piotr Nowak z Instytutu Matematycznego Polskiej Akademii Nauk, odczyt pt. *Sztywność grup – dowody wspierane komputerowo.* (s. 12)

11 czerwca 2019 r. – Dr Leszek Mellibruda z Active Business Mind

Psychologia biznesu, odczyt pt. *Autorytaryzm i populizm – dwa światy czy dwa oblicza tego samego?*

KONWERSATORIUM PW - SCIENTIA SUPREMA

W ramach Konwersatorium PW odbyły się również dwa odczyty z cyklu SCIENTIA SUPREMA czyli wykłady zainspirowane Nagrodą Nobla.

Pierwszy z nich miał miejsce 29 listopada 2018 r. Tym razem spotkanie było poświęcone tematyce, za którą została

↓ Prof. Marek Demiański podczas odczytu „Nagroda Nobla z fizyki 2019" na Wydziale Fizyki PW

przyznana w 2018 roku Nagroda Nobla w dziedzinie fizyki – przełomowe wynalazki w dziedzinie fizyki laserowej. Profesor Czesław Radzewicz z Wydziału Fizyki Uniwersytetu Warszawskiego wygłosił wykład pt. *W kierunku mocnych ultrakrótkich impulsów laserowych*.

Kolejny wykład w tym cyklu odbył się 14 listopada 2019. Profesor Marek Demiański z Uniwersytetu Warszawskiego wygłosił odczyt pt. *Nagroda Nobla z fizyki 2019*. Był on poświęcony pracy Jamesa Peeblesa, Michela Mayora i Didiera Queloza oraz przyznanej im nagrodzie za „ich wkład w nasze rozumienie ewolucji Wszechświata i miejsca Ziemi w kosmosie”. Wykład prof. M. Demiańskiego przedstawił główne osiągnięcia tegorocznych laureatów i przypomniał wczesne fazy ewolucji Wszechświata, proces powstawania lekkich pierwiastków i tworzenia się tła promieniowania reliktowego. Zostały omówione podstawowe metody poszukiwania egzoplanet i obecny stan wiedzy o układach planetarnych.

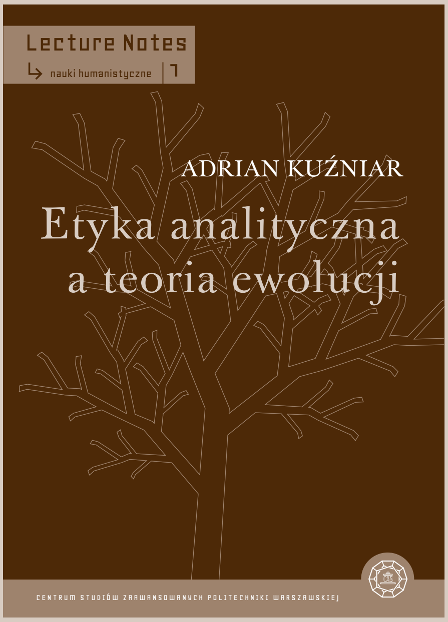
WYKŁADY UOD CSZ PW 2018/2019

W Uczelnianej Ofercie Dydaktycznej Centrum Studiów Zaawansowanych PW w roku akademickim 2018/2019 zorganizowano 6 wykładów podstawowych oraz 14 specjalnych, także w formie laboratoriów. W wykładach uczestniczyło łącznie ponad 800 osób, głównie doktorantów PW.

Słuchacze mieli możliwość wyboru spośród różnorodnej oferty tematycznej, zarówno kontynuacji jak i nowości:

1. *Równania różniczkowe zwyczajne*, prof. Wojciecha Domitrza
2. *Zaawansowane materiały*, prof. Małgorzaty Lewandowskiej
3. *Przetwarzanie i analiza danych w języku Python*, prof. Marka Gągolewskiego
4. *Podstawy sensorów chemicznych i biosensorów – budowa i zastosowania*, prof. Elżbiety Malinowskiej
5. *Wybrane Zagadnienia Termodynamiki Technicznej*, prof. Tomasza Wiśniewskiego
6. *Wstęp do analizy geometrycznej*, prof. Stanisława Janeczko
7. *Czy mogliśmy zrobić coś innego, niż uczyniliśmy? Analityczne wprowadzenie do filozoficznego sporu o wolność, determinizm i realność przyszłości,* dra hab. Adriana Kuźniara
8. *Wstęp do mechaniki kwantowej*, dra hab. Andrzeja Dragana
9. *Elementy współrzędnościowych systemów pomiarowych*, prof. Adama Woźniaka
10. *Laboratorium doskonalenia skuteczności i satysfakcji w relacjach z innymi ludźmi*, dra Leszka Mellibrudy (2 grupy)
11. *Wspomaganie Decyzji przy Wielorakości Celów w Praktyce Inżynierskiej – Wybrane Zagadnienia,* prof. Marianny Jacyny
12. *Konstrukcja Modeli Statystycznych z Pakietem R*, dr hab. Anny Dembińskiej
13. *Wnioskowanie Statystyczne z Pakietem R*, dr hab. Anny Dembińskiej
14. *Metody prowadzenia badań i statystycznej analizy wyników*, prof. Marka Dobosza
15. *Jak wydobyć potencjał twórczy grupy? Techniki pracy twórczej w grupie*, dra Bartłomieja Skowrona
16. *Wymykająca się osobowość*, dra Leszka Mellibrudy (2 grupy)
17. *Wprowadzenie do strukturalnej stabilności*, prof. Stanisława Janeczko
18. *Niezwykle szczególna teoria względności – wstęp*, dra hab. Andrzeja Dragana

*) Szczegółowy wykaz przedmiotów znajduje się na stronie www.konwersatorium.pw.edu.pl/oferta



WYDAWNICTWA

25 września 2019 gościem Centrum Studiów Zaawansowanych był dr Remi Lodth, edytor naukowy z wydawnictwa *Axel Springer SE*, którego częścią jest *Springer Nature*, wiodące na świecie wydawnictwo naukowe, edukacyjne i branżowe. Wizyta związana była z działalnością wydawniczą CSZ.

W grudniu 2019 roku ukazała się pierwsza książka w serii Lecture Notes – nauki humanistyczne – „Etyka analityczna a teoria ewolucji" autorstwa dra hab. Adriana Kuźniara z Uniwersytetu Warszawskiego.

Publikacje CSZ, w tym liczne książki naukowe, można nabyć w księgarniach Oficyny Wydawniczej PW w Gmachu Głównym i przy ul. Noakowskiego 18/20 w Warszawie.

*) Przegląd wszystkich dotychczasowych pozycji można odnaleźć pod adresem: www.csz.pw.edu.pl/Wydawnictwa.

DYSPUTY PITAGOREJSKIE

Dysputy stanowią formę rozmów, interakcji i spotkań inspirujących do dostrzegania nowych, ukrytych i zapomnianych aspektów rzeczywistości. Do uczestnictwa zapraszani są wybitni goście jako prowadzący oraz studenci PW, jako aktywni współuczestnicy dyskusji. 27 listopada 2018 odbyło się V spotkanie z cyklu „Dysput Pitagorejskich” pt. *Inteligencja naturalna i sztuczna – oczekiwania i prognozy*. Dysputę poprowadzili: prof. Stanisław Janeczko, prof. Piotr Przybyłowicz, dr hab. Andrzej Dragan, dr hab. inż. Przemysław Biecek oraz prof. Jerzy Rużyłło.

SEMINARIUM SPECJALISTYCZNE – KONWERSATORIUM PW

W semestrze zimowym Centrum Studiów Zaawansowanych we współpracy z Wydziałem Zarządzania oraz Centrum Informatyzacji Politechniki Warszawskiej zorganizowało i przeprowadziło kolejne seminaria z cyklu „Wyzwania modelowania inżynierskiego i biznesowego”:

16.10.2018 – *Transdyscyplinarne podejście do wybranych zagadnień zarządzania*, prof. Lech Bukowski z Wyższej Szkoły Biznesu w Dąbrowie Górniczej.

13.11.2018 – *Kreatywność a Industry 4.0. Czy maszyny mogą być kreatywne? Perspektywy i możliwości wykorzystania dorobku psychologii kreatywności na rzecz machine learning*, dr hab. Jan Fazlagić z Uniwersytetu Ekonomicznego w Poznaniu.

11.12.2018 – *Współczesna polityka pieniężna*, prof. dr hab. Eugeniusz

↑ Spotkania Otwartych Umysłów z udziałem Hassana Babikera, Andrzeja Dragana, Jana Fronka i Stanisława Janeczko

Gatnar z Uniwersytetu Ekonomicznego w Katowicach.

15.01.2019 – *Logistyka jako modelowy mix narzędzi zarządzania operacyjnego*, dr hab. Maciej Stajniak, Wyższa Szkoła Logistyki.

W semestrze letnim odbyły się następujące wykłady:

12.03.2019 – *Jacy jesteśmy? Przemiany postaw społeczeństwa polskiego w procesie demokratyzacji*, prof. dr hab. Iwona Jakubowska-Branicka z Uniwersytetu Warszawskiego.

9.04.2019 – *Podejmowanie decyzji a tryby i błędy myślenia*, dr hab. Beata Jamka.

14.05.2019 – *Idea przedsiębiorstw zwinnych*, prof. dr hab. inż. Stefan Trzcieliński z Politechniki Poznańskiej.

Ponadto, poza cyklem „Wyzwania modelowania inżynierskiego i biznesowego" odbyły się:

20 września 2018 – *Two problems connected with the Riemann hypothesis* – seminarium poprowadził prof. Don Bernard Zagier z Max Planck Institute for Mathematics w Bonn;

16 października 2018 – *Rigidity of Legendrian singularities* – seminarium poprowadził prof. Jun-Muk Hwang z Korea Institute for Advanced Study KIAS.

III KONFERENCJA Z FILOZOFII NAUKI I METOD FORMALNYCH W FILOZOFII

W dniach 13-14 grudnia 2018 r. odbyła się III konferencja z Filozofii Nauki i Metod Formalnych w Filozofii, współorganizowana przez Centrum Studiów Zaawansowanych wraz z Wydziałem Administracji i Nauk Społecznych PW oraz innymi instytucjami naukowymi.

WINTER WORKSHOP ON COMPLEX SYSTEMS

W dniach 4-8 lutego 2019 w Zakopanem i Krakowie odbyły się, pierwszy raz w Polsce, Zimowe Warsztaty Układów Złożonych (Winter Workshop on Complex Systems http://wwcs2019.org/) organizowane przez Wydział Fizyki Uniwersytetu Warszawskiego, przy wsparciu m.in. Wydziału Fizyki Politechniki Warszawskiej i Centrum Studiów Zaawansowanych PW. (s. 35)

SPOTKANIA OTWARTYCH UMYSŁÓW

28 lutego 2019 r. miało miejsce trzecie spotkanie z cyklu „Spotkania Otwartych Umysłów", na którym prowadzący: Hassan Babiker, Andrzej Dragan, Jan Fronk i Stanisław Janeczko, przy aktywnym współudziale słuchaczy omawiali następującą tematykę: ŻYCIE – wyższe i niższe formy inteligencji, POWSTAWANIE ŻYCIA – warunki konieczne, SYMULACJA ŻYCIA – procesy ewolucyjne.

SEMINARIUM CSZ ORAZ MINI PW

W dniach 17-18 maja 2019 roku odbyło się na Politechnice Warszawskiej międzynarodowe seminarium z matematyki dyskretnej współorganizowane przez Centrum Studiów Zaawansowanych i Wydział Matematyki i Nauk Informacyjnych Politechniki Warszawskiej.

Seminarium ma długą tradycję. Zostało zapoczątkowane w połowie lat 90. XX wieku przez prof. Michała Karońskiego z Uniwersytetu Adama Mickiewicza w Poznaniu i prof. Hansa Jürgena Prömela, wówczas z Humboldt University w Berlinie. Seminarium dotyczyło szerokiego zakresu tematycznego: ekstremalnej i probabilistycznej kombinatoryki, algorytmicznej matematyki dyskretnej oraz tematów pokrewnych. W ubiegłych latach seminarium było organizowane na przemian w Berlinie, Poznaniu i Hamburgu. Tym razem po raz pierwszy odbyło się w Warszawie.

W sesjach seminarium wzięli udział matematycy z Polski i Niemiec. Zaproszeni naukowcy wygłosili kilkanaście referatów.

IFIP NETWORKING 2019

W dniach 20-22 maja 2019 roku na Politechnice Warszawskiej odbyła się międzynarodowa konferencja IFIP NETWORKING 2019, przy współudziale Centrum Studiów Zaawansowanych, gdzie głównym prelegentem był profesor wizytujący CSZ, JJ Garcia Luna Aceves z University of California, Santa Cruz, USA.

SYMPOZJA CENTRUM STUDIÓW ZAAWANSOWANYCH

W 2019 roku Centrum zorganizowało dwa kolejne sympozja:

Przekraczanie granic struktur sieciowych – Sympozjum odbyło się w dniach 28-30 czerwca 2019 roku w Sandomierzu. Poświęcone było tematyce bezpieczeństwa w sieci, najnowszym osiągnięciom w tej dziedzinie, możliwości uczenia się maszyn, wykorzystaniu metod sztucznej inteligencji w praktyce oraz współczesnych naukach technicznych. (s.11)

Analiza geometryczna i zastosowania – sympozjum odbyło się 14-17.11.2019 w Ośrodku Badawczo-Konferencyjnym

↑ Prof. Andrzej Trautman, laureat Piątego Wyróżnienia „Kosmos Pitagorasa"

Instytutu Matematycznego PAN w Będlewie. Jego uczestnikami byli znakomici goście, wybitni naukowcy – często wieloletni współpracownicy Centrum. W ramach sympozjum wybrani goście wygłosili swoje odczyty. (s. 11)

PROFESOROWIE WIZYTUJĄCY Z EUROPEJSKIEGO TOWARZYSTWA MATEMATYCZNEGO

We wrześniu 2019 Centrum Studiów Zaawansowanych gościło członków Europejskiego Towarzystwa Matematycznego (EMS): prof. Volkera Mehrmanna, prezesa EMS; prof. Armena Sergeeva, wiceprezesa EMS, prof. Alice Fialowski oraz prof. Rolanda Duduchavę.

PIĄTE WYRÓŻNIENIE CENTRUM STUDIÓW ZAAWANSOWANYCH PW

3 grudnia 2019 roku po raz piąty zostało przyznane Wyróżnienie Centrum Studiów Zaawansowanych Politechniki Warszawskiej zwane „Kosmosem Pitagorasa", którego motto brzmi *Laus tibi, non tuleris qui vincula mente animoque* – „Chwała Ci za to, że nie pozwoliłeś nałożyć więzów na swój umysł i swego ducha". Wyróżnienie jest przyznawane na wniosek przewodniczącego Rady za szczególne zasługi dla środowiska naukowego. Za osiągnięcia w budowaniu efektywnych przestrzeni badawczych i przełamywanie granic poznania. Przy nominacji do wyróżnienia brane są pod uwagę cechy indywidualne kandydata, tj. osobowość, niezależność, skromność, oddanie, kreatywność, pokora, wykształcenie, energia. Statuetka wyróżnienia jest kryształowym dwunastościanem symbolizującym kryształową osobowość Mistrza, ponieważ jedynie Mistrz wie, jak wypełnić pustkę i nie jest więźniem materii, tylko Mistrz słyszy harmonię, wie skąd przychodzi, gdzie się znajduje i dokąd zmierza.

Tegorocznym laureatem wyróżnienia został profesor Andrzej Trautman, wybitny polski fizyk z Uniwersytetu Warszawskiego. Z tej okazji laureat wygłosił odczyt pt. „Moja droga od fal elektromagnetycznych na PW do fal grawitacyjnych na UW". Wręczenia wyróżnienia oraz pamiątkowej statuetki dokonał Prorektor Uczelni, prof. Rajumd Bacewicz wraz z dyrektorem Centrum, prof. Stanisławem Janeczko, w obecności wielu gości.

NAJBLIŻSZE PLANY

SYMPOZJUM W SANDOMIERZU

W maju 2020 r. planowane jest kolejne sympozjum Centrum Studiów Zaawansowanych PW, z udziałem profesorów z Politechniki Warszawskiej, Uniwersytetu Warszawskiego, Polskiej Akademii Nauk.

MASTER CLASS IN GEOMETRIC ANALYSIS

W czerwcu 2020 r. odbędzie się Master Class in Geometric Analysis z udziałem prof. Ying Ying Zhang z Uniwersytetu Tsinghua w Pekinie, przeznaczone dla polskich i chińskich studentów matematyki z Politechniki Warszawskiej i Uniwersytetu Tsinghua.

Jowita Krakowiecka

← (od lewej) Prof. Roland Duduchav, prof. Volker Mehrmann, prof. Alice Fialowski, prof. Stanisław Janeczko, prof. Armen Sergeev, podczas spotkania członków Europejskiego Towarzystwa Matematycznego (EMS)

„Przez miliony lat, ludzkość żyła tak jak zwierzęta. Potem stało się coś, co uwolniło siłę naszej wyobraźni. Nauczyliśmy się rozmawiać, nauczyliśmy się słuchać. Rozmowa umożliwiła przekazywanie pomysłów, dzięki czemu nauczyliśmy się wspólnie budować to, co niemożliwe. Największe osiągnięcia ludzkości tworzone są w rozmowie, a jej największe niepowodzenia są skutkiem braku rozmowy. Nie musi tak być. Nasze największe marzenia mogą stać się rzeczywistością. Z techniką, którą dysponujemy, możliwości są nieograniczone. Wszystko, co musimy zrobić, to upewnić się, że wciąż rozmawiamy."

Prof. Stephen Hawking (1942-2018)

ziemskim ekosystemie. Jedną z ważniejszych umiejętności *homo sapiens* stała się rozwinięta komunikacja werbalna, która finalnie doprowadziła do umiejętności rozmowy i przekazywania za jej pomocą myśli oraz emocji. Rozmowa jest aktem, w którym strony komunikujące się zarówno mówią, jak i słuchają, niekiedy dzięki rozmowie zmieniają swoje myśli, a także emocje. Przez tysiące lat ludzie na różny sposób próbowali zrozumieć otaczający świat tworząc różne teorie na temat relacji pomiędzy zjawiskami głównie fizycznymi. Empedokles z Akragas wyróżnił cztery żywioły: ziemię, wodę, powietrze i ogień. W kulturze japońskiej pojawia się pustka oznaczająca też niebo, w kulturze chińskiej znika powietrze, a pojawia się drewno i metal. Z kolei w hinduizmie jako żywioł pojawia się dźwięk. Teorie żywiołów mają swoje odniesienie do przyrody, którą w dużym stopniu była eksplorowana przez ludzi poprzez obserwację, rzadziej przez przeprowadzenie doświadczenia. Żywioły mają charakter fizyczny, namacalny, są w dużym stopniu pozbawione kontroli, chociaż nie ulega wątpliwości, że ludzie posiadają potrzebę okiełznania ich.

Celem artykułu jest przedstawienie trzech tez:

1. kulturowej: cyberprzestrzeń jest zaawansowanym tworem techniczno-kulturowym – jest realizacją marzeń wielu twórców, wynalazców i inżynierów,
2. technicznej: bezpieczeństwo i cyberprzestrzeń to nierozłączne elementy (stąd: cyberbezpieczeństwo),
3. paranoicznej: pełne bezpieczeństwo, jeśli jest osiągalne, nie jest stanem stałym (stąd tytułowe: cyber(nie)bezpieczeństwo).

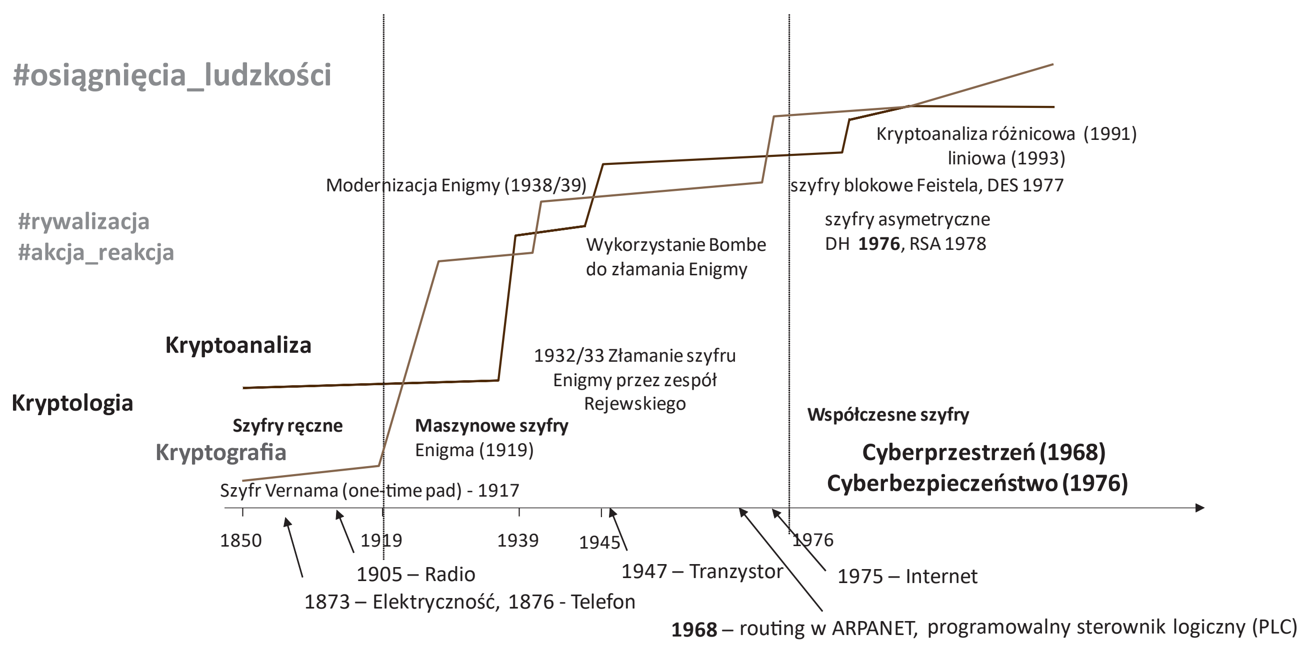
Artykuł ma charakter bardziej poglądowy, niż przeglądowy i składa się z pięciu rozdziałów, którego pierwszym jest niniejsze krótkie wprowadzenie. Drugi rozdział dotyczy genezy cyberprzestrzeni i jej bezpieczeństwa. Trzeci opisuje przebieg rozmowy w cyberprzestrzeni i związane z tym zagrożenia. Czwarty dotyczy zmienności i ryzyka wpływających na niebezpieczne związki, które stają się kluczowe przy projektowaniu systemów zabezpieczeń. Ostatni jest zwięzłym podsumowaniem niniejszej pracy.

2. Geneza cyberprzestrzeni i jej bezpieczeństwa[1]

Cyberprzestrzeń pojawia się w historii ludzkości w drugiej połowie XX wieku. Jest wynikiem procesu tworzenia wynalazków i różnych teorii, z których wiele z nich ułatwiło komunikację. Zgodnie z [1] cyberprzestrzeń (ang. *cyberspace*) jest z jednej strony zbiorem technik cyfrowych służących do wymiany informacji, ale z drugiej nowego typu przestrzenią społeczną częściowo wirtualną, która może być bytem całkowicie odseparowanym od fizycznego. Z trzeciej jest środowiskiem, w którym następuje współczesna komunikacja za pomocą sieci teleinformatycznej. W psychologii ewolucyjnej zaproponowano nowe pojęcie *homo cyberus* [5] skojarzone z cybersocjalizacją. Pojęcie

[1] *Wykorzystano fragment opracowania autorskiego, które wchodzi w skład [1]*

↓ Rysunek 1. Cyberprzestrzeń: początki, w tle kryptologii nowożytnej



„CYBERSPACE" [3] (pisane wersalikami) pojawia się w cyklu kolaży tworzonych w latach 1968-1970 przez duńską artystkę Susanne Ussing we współpracy z duńskim architektem Karstenem Hoffem. Dekadę później termin pojawia się w literaturze – używa go twórca cyberpunku William Gibson [4]. O ile kolaże Ussing stanowią ekspresyjne wizje ludzi wtopionych w cybernetyczne formy, co raczej należy uznać za neutralne w odbiorze, o tyle cyberpunk, jako odmiana fantastyki naukowej, skupia się przede wszystkim na negatywnych stronach cyfryzacji świata i zetknięcia *homo sapiens* z maszynami. Tak małymi krokami objawia się nowy żywioł bez znamion cech fizycznych, w którym nie obowiązują prawa fizyki takie jak chociażby zasady dynamiki Newtona, w tym prawo ciążenia.

Bezpieczeństwo pojawia się w zasadzie od razu z cyberprzestrzenią jako jego immanentna część. Najpierw jest rozumiane dość naiwnie na poziomie fizycznym, jako odseparowanie sygnałów istotnych (np. tajnych) od pozostałych wręcz na poziomie duktów i okablowania (tzw. *air gap*). To oczywiście, już na pierwszy rzut oka, wydaje się niewystarczające i stąd do ochrony komunikacji zostaje zaprzątnięta kryptologia, w szczególności kryptografia. Zatem od strony technicznej powstanie cyberprzestrzeni i pojawienie się w niej (cyber)bezpieczeństwa można spróbować skorelować z rozwojem kryptologii nowożytnej (Rysunek 1). Z kolei sam rozwój kryptologii i w konsekwencji cyberbezpieczeństwa jest ściśle powiązany z prowadzonymi działaniami wojennymi. Kryptologia jest nauką, w której występują dwie siły: syntetyzująca, budująca (kryptografia) i analizująca, wręcz niszcząca (kryptoanaliza). W historii kryptologii te siły wzajemnie napędzają się i konkurują ze sobą – metody kryptograficzne muszą być odporne na znane ataki kryptoanalityczne. Powstaje swojego rodzaju wyścig zbrojeń, w którym algorytmy są ulepszane, a głównym miernikiem ich siły jest czas ich ochrony, wynikający w dużym stopniu z długości klucza. Na tę chwilę kresem tego wyścigu od strony kryptoanalizy jest wizja pojawienia się komputera kwantowego [6], a od strony kryptografii – algorytmy postkwantowe [7].

Koniec I wojny światowej to cezura końcowa epoki szyfrów ręcznych. Ponad wiek temu, w 1917 roku pojawia się algorytm Vernama, będący szyfrem idealnym, którego istotą jest wykonywanie dodawania modulo 2 wiadomości z losowym kluczem o długości tej samej co wiadomość (stąd określenie: szyfr z kluczem jednorazowym). Wejście w epokę szyfrów maszynowych, która trwa do połowy zimnej wojny, symbolizuje powstanie w Niemczech pierwszej wersji elektromechanicznej maszyny szyfrującej Enigma [8]. Marian Rejewski, Jerzy Różycki i Henryk Zygalski doprowadzają do złamania Enigmy, czego efektem jest modernizacja Enigmy (dodanie kolejnych rotorów) tuż przed wybuchem II wojny światowej. W trakcie wojny ulepszona Enigma zostaje złamana z wykorzystaniem Bombe [9] – elektromechanicznej maszyny stworzonej w Wielkiej Brytanii przez Alana Turinga na podstawie projektu i prototypu grupy Rejewskiego tzw. bomby kryptologicznej. W 1947 roku zostaje skonstruowany pierwszy tranzystor ostrzowy przez Johna Bardeena oraz Waltera Housera Brattaina. Staje się on punktem wyjścia do stworzenia współczesnych maszyn obliczeniowych, co następuje w ciągu kolejnej dekady. Pod koniec epoki szyfrów maszynowych w 1968 pojawia się *routing* (trasowanie) w sieci ARPANET [10], a także programowalny sterownik logiczny (*Programmable Logic Controller* – PLC) [11]. Rok 1968 możemy uznać za datę inicjacji cyberprzestrzeni. Sieć Internet w swojej pierwotnej formie debiutuje mniej więcej siedem lat później [12].

Cezura początkowa dla epoki współczesnych szyfrów następuje rok później po pojawieniu się Internetu – w 1976 zostaje opublikowany przez Witfielda Diffiego oraz Martina Hellmana [13] pierwszy algorytm wymiany klucza współdzielonego oparty na algorytmach asymetrycznych wykorzystujący złożoność obliczenia logarytmów dyskretnych. Zatem rok 1976 możemy uznać za symboliczną datę inicjacji cyberbezpieczeństwa. Rok później zostaje zaakceptowany jako standard federalny w USA [14] blokowy algorytm symetryczny DES (*Data Encryption Standard*), który staje się w 1981 roku standardem sektora prywatnego na kolejne dwie dekady. W 1978 roku Ron Rivest, Adi Shamir oraz Leonard Adleman publikują pracę [15] dotyczącą pierwszego kryptosystemu wykorzystującego klucz publiczny i problem złożoności faktoryzacji dużych liczb, nazwanego od liter ich nazwisk RSA.

Współczesnym motorem cyberbezpieczeństwa nadal pozostają działania zbrojne. W USA po okresie zimnej wojny, podczas której rozwinęła się Dolina Krzemowa [16], głównie przez ogromne inwestycje kapitałowe, zamachy z 11 września 2001 otworzyły etap walki z ogólnie pojętym terroryzmem. Z kolei drugi kraj przodujący na świecie w cyberbezpieczeństwie – Izrael, który po zakończeniu drugiej wojny światowej zaangażował się w długotrwały konflikt z Palestyną, przyjął podobny model jak USA podczas zimnej wojny – inwestycje w branżę zbrojeniową, która formalnie od dekady jest ściśle powiązana z cyberprzestrzenią. Przykładowo w USA Cyber Command [17] powstał w 2009 roku i uzupełnia działania sił zbrojnych USA dotychczas powiązanych z tradycyjnymi żywiołami (US Air Force – powietrze, US Army – ziemia, US Navy – woda) w cyberprzestrzeni.

Cyberprzestrzeń jest także obecna w przemyśle. Patrząc na wprowadzoną kilka lat temu klasyfikację [18]:

- Przemysł 1.0: wiek pary,
- Przemysł 2.0: wiek elektryczności,
- Przemysł 3.0: wiek komputerów,
- Przemysł 4.0: wiek zanikania bariery ludzie-maszyny,

cyberprzestrzeń pojawia się w tzw. wieku komputerów, gdzie masowa produkcja zostaje wsparta przez maszyny. Historycznie masowa produkcja powstała przez dodanie do przemysłu elektryczności, a z kolei sam przemysł przez wsparcie ręcznej produkcji maszynami parowymi. Wiek komputerów i sterowanie nimi za pomocą systemów SCADA (*Supervisory Control and Data Acquisition*) nie byłby możliwy, bez wspomnianego wcześniej, opracowania sterownika logicznego PLC (*Programmable Logic Controller*) w 1968 [11]. Za ojca tego wynalazku uznaje się Dicka Morley amerykańskiego inżyniera i wynalazcę. Urządzenie to pozwalało na fizyczne sterowanie maszyną i realizowanie konkretnego algorytmu niezbędnego do procesu produkcji. Natomiast przemysł 4.0 bazuje w dużym stopniu na dołączeniu systemów SCADA do Internetu (co technicznie nie jest problemem od dawna) i na wykorzystanie rozwiązań z obszaru przemysłowego Internetu rzeczy (*Industrial Internet of Things* – [19]). W tym przypadku pojawia się ogromna liczba danych wymagających zabezpieczenia.

Reasumując za datę powstania cyberprzestrzeni, zarówno w sensie kulturowym jak i technicznym, można przyjąć rok 1968. Cyberbezpieczeństwo pojawia się 8 lat później za sprawą

↑ Rysunek 2. a) Rozmowa[2]; b) Cyfryzacja mowy – pojawiają się dane[2]; c) Podsłuch[2]; d) Modyfikacja2; e) Podszycie[3]; f) Wyparcie się[2]

pierwszych algorytmów wykorzystujących kryptografię asymetryczną.

3. Rozmowa w cyberprzestrzeni i związane z nią zagrożenia

Wróćmy do wątku rozmowy (Rysunek 2a[2]). Zanim mowa znajdzie się w cyberprzestrzeni musi zostać poddana cyfryzacji, czyli zmiany postaci analogowej na cyfrową (Rysunek 2b[2]). Cyfryzacja sygnału mowy składa się z dwóch procesów: próbkowania i kwantyzacji. Zgodnie z twierdzeniem Nyquista dot. próbkowania, częstotliwość próbkowania jest dwukrotnością próbkowanego kanału i przykładowo: 8 kHz – dla klasycznej telefonii stacjonarnej, 16 kHz albo 22,050 kHz – dla średniej jakości (radio UKF) oraz 44,1 kHz albo 48 kHz – dla jakości płyty kompaktowej (HiFi). Kwantyzacja jest procesem, w którym próbki są konwertowane na ustalone wartości liczbowe. Tak powstają dane – będące w omawianym przypadku przekształceniem mowy na postać binarną.

Dane rozpoczynają swoją podróż w cyberprzestrzeni: mogą być przesyłane (są danymi „w ruchu"), magazynowane lub przetwarzane. W każdym momencie swojej podróży mogą być podsłuchane (Rysunek 2c[2]), bądź zmodyfikowane (Rysunek 2d[2]). Istnieje możliwość podszycia się (Rysunek 2e[3]) pod wysyłającego dane (pod odbierającego też), a także wyparcie się faktu wysłania lub odebrania danych (Rysunek 2f[2]). Są to podstawowe zagrożenia, które dotyczą danych. Podsłuch jest działaniem pasywnym – nie wpływa na strukturę danych, pozostałe trzy zagrożenia wymagają aktywności i w zależności od źródła danych, mniej lub więcej zaangażowania. Najprościej operować na danych stworzonych od początku do końca w cyberprzestrzeni, a nie, tak w jak przykładzie rozmowy, biometrycznych – wtedy np. modyfikacja jest o wiele łatwiejsza do przeprowadzenia.

Dane jako takie, bez żadnego kontekstu, są wyłącznie binarnym zbiorem, są literami o nieuporządkowanej strukturze. Nabierają wartości dopiero, jeśli stanowią informacje [22], które można porównać już do słów. Możemy mówić zarówno o ochronie danych, jak i o ochronie informacji. Obecnie w czasach dostępności ogromnych przepustowości sieciowych i mocy obliczeniowej istnieje pokusa na to, żeby chronić wszystkie dane w zasadzie bezmyślnie. Wydzielenie informacji ważnych i dopasowanie do nich odpowiednich metod ochrony jest jak najbardziej pożądane i zasadne. Wynika to chociażby z tego, że „czas życia" poszczególnych

[2] *Wykorzystano zdjęcie z [20]*

[3] *Wykorzystano zdjęcie z [21]*

informacji, a więc potencjalna potrzeba ich ochrony jest różna np. położenie samolotów wielozadaniowych General Dynamics F-16 Fighting Falcon nad Polską to poziom kilkudziesięciu minut, a tożsamość szpiegów to minimum pół wieku.

Na podstawie informacji jesteśmy w stanie budować wiedzę, która w kontekście cyberprzestrzeni jest magazynowaniem istotnych informacji, poszukiwaniem relacji i wyciąganiem wniosków. Wiedzę możemy porównać do zdań albo wręcz do myśli. Na końcu łańcucha przetwarzania danych możemy umieścić mądrość, jako umiejętność podejmowania właściwych (cokolwiek to znaczy!) decyzji na podstawie zdobytej wiedzy.

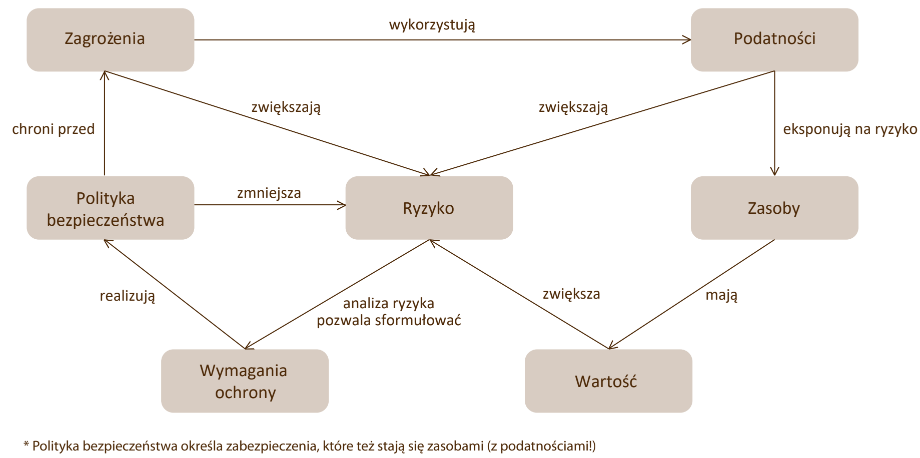
W kontekście wspomnianych zagrożeń (podsłuch, modyfikacja, podszycie się, wyparcie się) wprowadza się podstawowe usługi cyberbezpieczeństwa [23]: poufność, integralność, uwierzytelnienie i niezaprzeczalność. Uzupełnia się ten zestaw o usługę kontroli dostępu, dzięki której można sterować prawami korzystania z zasobów.

Usługi ochrony informacji są budowane za pomocą różnych mechanizmów, jednym z nich jest szyfrowanie leżące w obszarze kryptografii. Jak wspomniałem w poprzednim rozdziale głównym miernikiem siły algorytmów kryptograficznych jest czas ich ochrony, wynikający w dużym stopniu z długości klucza. Przy klasycznym modelu rozwoju technik obliczeniowych można dokonać predykcji szansy na złamanie chronionych informacji (także bazując na prawie Moore'a) i w ten sposób oferować ochronę na ustalony okres. Pojawiają się dwie ważne kwestie: pierwsza to umiejętne zarządzanie kluczami, w skład którego wchodzi ich generacja, dystrybucja, przechowywanie, a także zniszczenie po czasie użycia. Drugie to odporność danego algorytmu na znane ataki kryptoanalityczne, a także konstrukcja odporna na tzw. tylne furtki, czyli sposoby szybszego przeliczenia pewnych własności z wykorzystaniem słabych cech algorytmu [24].

4. Niebezpieczne związki: projektowanie systemów zabezpieczeń

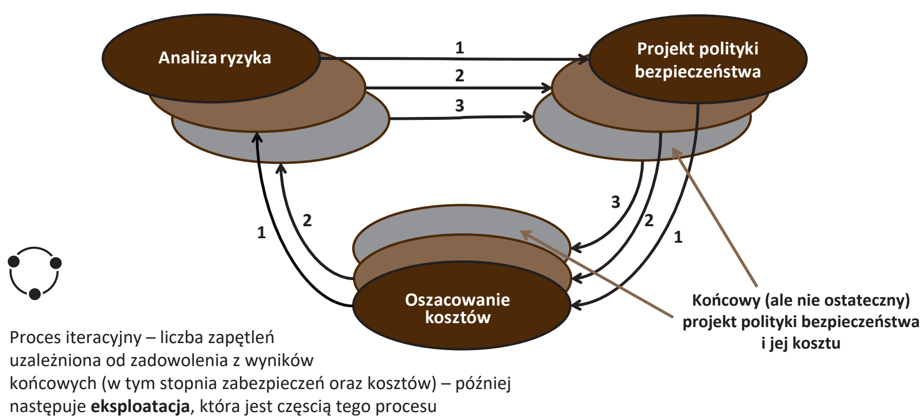
Zagrożenia są pierwotną przyczyną działań w obszarze cyberbezpieczeństwa. Zagrożenia (Rysunek 3) wykorzystują podatności, które pojawiają się na etapie projektowania, wdrażania, bądź konfigurowania systemów teleinformatycznych. Podatności (podobnie jak zagrożenia) zwiększają ryzyko, które jest wskaźnikiem stanu lub zdarzenia, które może prowadzić do strat. Podatności eksponują na ryzyko zasoby, które mają wartość (nie tylko ekonomiczną). Im większa wartość tym potencjalnie większe ryzyko. Polityka bezpieczeństwa definiuje zabezpieczenia na poziomie organizacyjnym oraz technicznym i dzięki temu zmniejsza ryzyko i chroni przed zagrożeniami. Z kolei wymagania ochrony realizują politykę bezpieczeństwa, a dzięki kluczowej przy projektowaniu zabezpieczeń analizie ryzyka [25], dają się sformułować. Warto zauważyć, że polityka bezpieczeństwa wprowadza nowe zasoby wraz z nowymi, właściwymi dla nich, podatnościami.

Analiza ryzyka, polegająca na przewidzeniu negatywnych skutków działań oraz zjawisk i odpowiednim obniżeniu potencjalnych strat wynikających z takich sytuacji jest procesem (Rysunek 4), który, jak wspomniano wcześniej, skutkuje projektem polityki bezpieczeństwa. Tak przedstawiony projekt podlega oszacowaniu kosztów. Jest to proces iteracyjny – przyjęcie akceptowalnego poziomu ryzyka przy akceptowalnych kosztach daje punkt wyjścia do przyjęcia polityki bezpieczeństwa i jednocześnie ciągły – w trakcie eksploatacji systemów zabezpieczeń powinien przebiegać okresowo.



↑ Rysunek 3. Niebezpieczne związki

System teleinformatyczny możemy obserwować poprzez składowe obiekty i relacje między nimi [26]. W każdym obiekcie możemy obserwować cechy, które mogą mieć charakter stały, ale część z nich może podlegać zmianie. Zmiana jest związana z przejściem cechy w inną. Obserwacja cechy jest ważna, wtedy tylko, kiedy można z informacją o zmianie tej cechy coś zrobić. Nie każda obserwacja zatem jest wartościowa z praktycznego punktu widzenia. W obiekcie można wyróżnić wiele cech do obserwacji, jednak w wielu przypadkach nie posiadamy pełnej wiedzy o wszystkich cechach. Jednocześnie pewne cechy mogą niespodziewanie pojawiać się, a inne z kolei znikać. Z drugiej strony część z cech nie jest istotna dla obserwacji zmiany bądź co gorsza obserwator nie ma świadomości ich istotności. Po wyróżnieniu konkretnej cechy, obserwacji podlega zmiana własności cechy w porównaniu do różnych miar, często



↑ Rysunek 4. Projektowanie systemów bezpieczeństwa

→

wyliczanych na podstawie poprzednich okresów obserwacji. W przypadku sieci teleinformatycznych obiektem jest przeważnie protokół telekomunikacyjny. Dla każdego protokołu budowany jest osobny model wyróżniający istotne cechy np. połączeniowość, mechanizmy retransmisji, obsługę opóźnionych lub uszkodzonych jednostek danych. Badając każdy z tych aspektów, analizie podlegają parametry sieciowe skojarzone z daną cechą przy mechanizmach retransmisji (liczniki czasu typu *timeout*, wielkość okna retransmisji, stopa błędów). Dopatrzenie się różnicy w profilu protokołu jest oznaką zmiany i z dużym prawdopodobieństwem oznacza atak sieciowy. Budowanie wiedzy na podstawie informacji, które są jednoznacznymi wskazaniami na atak, ułatwia tworzenie odpowiednich sygnatur, natomiast zjawiska, które są w jakimś stopniu nierozpoznane mają status anomalii. W eksploatacji systemów bezpieczeństwa zagadnienie fałszywych alarmów wpływa na rzetelność systemów wykrywających anomalie, w szczególności błędy drugiego rodzaju (*false positives*) prowadzą do nierozpoznawania ataków. Błędy pierwszego rodzaju (*false nagatives*) konsumują zasoby, ale nie są krytyczne dla bezpieczeństwa systemów (coś co nie jest atakiem w rzeczywistości, jest traktowane jako atak).

5. Podsumowanie

Cyberprzestrzeń jest zaawansowanym wytworem ludzkiej wyobraźni nie tylko w sensie technicznym, ale także w społeczno-kulturowym. Dzięki technice stworzona całkowicie w umyśle człowieka wizja stała się rzeczywistością – jest to przykład spełnienia marzeń.

Cyberprzestrzeń przenika się ze światem fizycznym, chociaż może być całkowicie wirtualna. W swojej naturze jest nierozłączna z bezpieczeństwem, które jest zmienne i nigdy nie jest pełne stąd tytułowe „cyber(nie) bezpieczeństwo".

Ludzie (w tym autor początkowego cytatu) od dawna marzyli o eksploracji kosmosu, a być może także o przeniesieniu życia na obcą planetę. Tymczasem mamy cyberprzestrzeń, być może będziemy w stanie replikować się w niej i dzięki temu staniemy się nieśmiertelni i w pełni szczęśliwi. Motyw ten jest znany z fantastyki naukowej (np. z filmu „Lucy" w reżyserii i wg scenariusza Luca Bessona [27]) i być może wymaga ponownej rewizji.

LITERATURA

[1] J. Woźniak, A. Bęben, J. Mongay Batalla, M. Natkaniec, Z. Piotrowski, K. Szczypiorski, K.Wesołowski – *Tendencje w rozwoju polskiej i światowej telekomunikacji i teleinformatyki* (w przygotowaniu), 2020

[2] https://en.wikipedia.org/wiki/Cyberspace (data pobrania: 21.10.2019)

[3] https://kunstkritikk.com/the-reinvention-of-cyberspace/ (data pobrania: 21.10.2019)

[4] https://en.wikipedia.org/wiki/William_Gibson (data pobrania: 21.10.2019)

[5] Pleshakov V.A. - *Human cyber socialization: from Homo Sapiens'a to Homo Cyberus'a*: monograph. M.: Publishing house of the Moscow State Pedagogical University "Prometey", 2012.221

[6] Frank Arute, Kunal Arya, Ryan Babbush, et al. – *Dave Bacon Quantum supremacy using a programmable superconducting processor*. Nature. Volume 574, pages 505–510 (23 October 2019)

[7] Daniel J. Bernstein, Tanja Lange – *Post-quantum cryptography*. Nature, Volume 549, pages 188-194 (14 September 2017)

[8] https://en.wikipedia.org/wiki/Enigma_machine (data pobrania: 21.10.2019)

[9] https://www.cryptomuseum.com/crypto/bombe/ (data pobrania: 21.10.2019)

[10] https://www.internetsociety.org/internet/history-internet/brief-history-internet/ (data pobrania: 21.10.2019)

[11] https://library.automationdirect.com/history-of-the-plc/ (data pobrania: 21.10.2019)

[12] https://tools.ietf.org/html/rfc675 (data pobrania: 21.10.2019)

[13] W. Diffie, M. Hellman – *New directions in cryptography*, IEEE Transactions on Information Theory, Volume 22 Issue 6, Nov. 1976, pp. 644-654

[14] Federal Information Processing Standard (FIPS) 46: Data Encryption Standard (DES), January 1977, https://csrc.nist.gov/publications/detail/fips/46/archive/1977-01-31 (data pobrania: 21.10.2019)

[15] R. L. Rivest, A. Shamir, L. Adleman – *A method for obtaining digital signatures and public-key cryptosystems*, Communications of the ACM, Volume 21 Issue 2, Feb. 1978, pp. 120-126

[16] https://en.wikipedia.org/wiki/Silicon_Valley (data pobrania: 21.10.2019)

[17] https://www.arcyber.army.mil (data pobrania: 21.10.2019)

[18] https://www.bmbf.de/de/zukunftsprojekt-industrie-4-0-848.html (data pobrania: 21.10.2019)

[19] H. Boyes, B. Hallaq, J. Cunningham, T. Watson – *The industrial internet of things (IIoT): An analysis framework*, Computers in Industry, Volume 10 1, 2018, pp. 1-12

[20] https://glavcom.ua/scotch/showbiz/kultovaya-trilogiya-matrica-oficialno-poluchit-perezapusk-483980/g358176.html (data pobrania: 21.10.2019)

[21] https://www.cinemablend.com/new/Matrix-Fan-Theory-Puts-Agent-Smith-One-It-Kind-Works-121077.html (data pobrania: 21.10.2019)

[22] B. Stefanowicz – *Informacja. Wiedza. Mądrość. Sześćdziesiąty* szósty tom Biblioteki Wiadomości Statystycznych, Główny Urząd Statystyczny, 2013

[23] ISO 7498-2:1989 – Information processing systems – Open Systems Interconnection – Basic Reference Model – Part 2: Security Architecture

[24] N. Perlroth, J. Larson, S. Shane – *N.S.A. Able to Foil Basic Safeguards of Privacy on Web.* 05.09.2013, New York Times – https://www.nytimes.com/2013/09/06/us/nsa-foils-much-internet-encryption.html (data pobrania: 21.10.2019)

[25] ISO 31000:2018 – Risk management – Principles and guidelines

[26] J. Bieniasz, M. Stępkowska, A. Janicki, K.Szczypiorski – *Mobile Agents for Detecting Network Attacks Using Timing Covert Channels*, Journal of Universal Computer Science, Volume 25 Issue 9, Sep. 2019, pp. 1109-1130

[27] Lucy (2014), https://www.imdb.com/title/tt2872732/ (data pobrania: 21.10.2019)

Artykuł na podstawie wykładu inauguracyjnego roku akademickiego 2019/2020 na Wydziale Elektroniki i Technik Informacyjnych PW wygłoszonego 1 października 2019 w Dużej Auli w Gmachu Głównym PW.

{**Krzysztof Szczypiorski** - jest profesorem Politechniki Warszawskiej na Wydziale Elektroniki i Technik Informacyjnych. Jest założycielem Zakładu Cyberbezpieczeństwa na macierzystym Wydziale, którego jest kierownikiem od 2015 roku. Ukończył studia na Politechnice Warszawskiej w 1997 roku, a w kolejnych latach uzyskał stopień doktora i doktora habilitowanego w dziedzinie telekomunikacji ze specjalizacją w zakresie bezpieczeństwa informacji. Ukończył również podyplomowe studia na SWPS w Warszawie oraz Hass School of Business na Uniwersytecie Kalifornia w USA. Jest współzałożycielem i kierownikiem nowego kierunku studiów na Wydziale Elektroniki i Technik Informacyjnych - Cyberbezpieczeństwo. Wizytował m.in. takie uczelnie jak: George Mason University, Fairfax, Virginia, USA (2014), Luxembourg Institute of Science & Technology, Belval Innovation Campus, Esch-sur-Alzette, Luksemburg (2015), University of California, Berkeley, USA (2013) i ponad 50 innych miejsc na krótkoterminowe pobyty naukowe. Od ponad 25 lat Krzysztof Szczypiorski jest niezależnym konsultantem w dziedzinie cyberbezpieczeństwa, telekomunikacji i informatyki dla wielu podmiotów, w tym: Cisco Systems, Hewlett-Packard, Ministerstwa Finansów (Polska), Biura Bezpieczeństwa Narodowego (Polska), Oracle, Orange, Parlamentu Rzeczypospolitej Polskiej, Polskiej Grupy Energetycznej, PwC, T-Mobile Polska. Jest autorem lub współautorem ponad 200 artykułów i ponad 70 zaproszonych rozmów, a także 3 zgłoszeń patentowych (jedno z nich jest przyznane).}

PRZEKRACZANIE GRANIC STRUKTUR SIECIOWYCH

Sympozjum Centrum Studiów Zaawansowanych PW

↑ Uczestnicy sympozjum w Sandomierzu „Przekraczanie granic struktur sieciowych"

W dniach 28-30.06.2019 w Sandomierzu odbyło się sympozjum Centrum Studiów Zaawansowanych „Przekraczanie granic struktur sieciowych".

Sympozjum poświęcone było tematyce bezpieczeństwa w sieci, najnowszym osiągnięciom w tej dziedzinie, możliwości uczenia się maszyn, wykorzystaniu metod sztucznej inteligencji w praktyce i współczesnych naukach technicznych. Poruszone zostały zagadnienia dotyczące możliwości i nadziei, ale również obaw i zagrożeń, a także etycznych dylematów związanych z obecnym dynamicznym rozwojem robotyki i sztucznej inteligencji oraz perspektyw rozwoju tych dziedzin.

Warsztaty Naukowe Centrum Studiów Zaawansowanych, których częścią jest to sympozjum, są uzupełnieniem oferty dydaktycznej Centrum. Ich podstawowym celem jest przełamanie barier utrudniających integrację ludzi nauki, wynikających z podziałów strukturalnych i pokoleniowych. Organizowane w ten sposób spotkania umożliwiają wymianę doświadczeń i nawiązanie współpracy naukowo-badawczej między uczestnikami reprezentującymi często różne dziedziny nauki i etapy kariery naukowej. Inspirujące dyskusje, które towarzyszą warsztatom, przyczyniają się natomiast do poszerzenia horyzontów naukowych specjalistów biorących udział w spotkaniu i mogą być początkiem współpracy.

Jowita Krakowiecka

ANALIZA GEOMETRYCZNA I ZASTOSOWANIA

Sympozjum Centrum Studiów Zaawansowanych PW

W dniach 14-17.11.2019 w Ośrodku Badawczo-Konferencyjnym Instytutu Matematycznego PAN w Będlewie odbyło się sympozjum pt. „Analiza geometryczna i zastosowania".

Sympozjum zorganizowane zostało przez Centrum Studiów Zaawansowanych PW, pod auspicjami komitetu naukowego w składzie: prof. Wojciech Domitrz – Dziekan Wydziału Matematyki i Nauk Informacyjnych PW; prof. Stanisław Janeczko – Dyrektor CSZ PW, Wydział Matematyki i Nauk Informacyjnych PW; prof. Wojciech Kucharz – Instytut Matematyki Uniwersytet Jagielloński; prof. Marek Kuś – Centrum Fizyki Teoretycznej PAN, Międzynarodowe Centrum Ontologii Formalnej, Wydział Administracji i Nauk Społecznych PW; prof. Piotr Przybyłowicz – Zastępca Dyrektora CSZ PW, Wydział Samochodów i Maszyn Roboczych PW.

Jowita Krakowiecka

↑ Uczestnicy sympozjum w Będlewie „Analiza geometryczna i zastosowania"

Sztywne grupy

Piotr W. Nowak z Instytutu Matematycznego Polskiej Akademii Nauk na postawie swojego odczytu 23 maja 2019 r. w ramach Konwersatorium Politechniki Warszawskiej

Mając grupę G spróbujmy zrealizować ją jako grupę izometrii przestrzeni Euklidesowej, skończenie lub nieskończenie wymiarowej. Innymi słowy, weźmy dowolny homomorfizm

$$\varphi\colon G \to Isom(H),$$

gdzie H jest przestrzenią Hilberta a $Isom(H)$ jej grupą izometrii, czyli przekształceń H samej w siebie zachowujących odległość. Każdy element

$$A \in Isom(H)$$

jest przekształceniem afinicznym, czyli składa się z części liniowej i przesunięcia: jest postaci

$$Av = Uv + b,$$

gdzie U jest przekształceniem unitarnym, czyli izometrią zachowującą iloczyn skalarny przestrzeni H, a

$$b \in H$$

wektorem przesunięcia.

Powiemy, że grupa G ma własność (T) jeśli obraz

$$\varphi(G)$$

dla dowolnego φ ma punkt stały: istnieje $v \in H$ takie, że

$$\varphi(g)v = v$$

dla każdego $g \in G$.

Istnienie takiego punktu stałego dla każdego homomorfizmu φ jest dużym ograniczeniem i oznacza, że grupa G nie daje się realizować w ciekawy sposób jako grupa izometrii przestrzeni Hilberta. Własność tą często nazywa się sztywnością grupy.

Sztywne grupy mają szereg zastosowań. Ich skończone ilorazy pozwalają konstruować rodziny nieskończonych grafów, tzw. ekspanderów, posiadających dość pozornie sprzeczne własności: mają jednostajnie ograniczony stopień wierzchołków, ale jednocześnie są bardzo silnie spójne – aby je rozspójnić należy wyjąć relatywnie dużą ilość krawędzi. Grafy takie używane są przy projektowaniu alogrytmów streamingujących, które odpowiedzialne są za wyszukiwanie trendów w mediach społecznościowych, np. na Twitterze [9]. Grupy z własnością (T) są też głównym źródłem kontrprzykładów na niektóre wersje hipotezy Bauma-Connesa. Hipoteza ta jest niezwykle udaną próbą uogólnienia klasycznego twierdzenia o indeksie, które w latach 60. XX wieku udowodnili M.F. Atiyah i I. Singer.

Przyjrzyjmy się zatem bliżej, które grupy są sztywne, a które nie. Dla przykładu, grupa $\mathbb{Z}$ nie posiada zdefiniowanej powyżej własności (T). Istotnie, rozważmy działanie grupy $\mathbb{Z}$ na przestrzeni $\mathbb{R}$ zdefiniowane w następujący sposób:

$$n \cdot x = x + n,$$

gdzie

$$n \in \mathbb{Z},\, x \in \mathbb{R}.$$

Łatwo sprawdzić, że jest to działanie przez izometrie, jednak w oczywisty sposób nie posiada ono punktu stałego, ponieważ dowolny $x \in \mathbb{R}$ jest przesuwany przez kolejne liczby całkowite w nieskończoność.

Zauważmy, że własność (T) przenosi się na ilorazy: jeśli grupa G ma własność (T) to dowolny iloraz G/N, gdzie N jest podgrupą normalną, również posiada (T). Istotnie, w tej sytuacji mając dane działanie grupy G/N poprzez izometrie na przestrzeni H możemy rozpatrywać je jako działanie G, poprzez homomorfizm ilorazowy $G \to G/N$. Własność (T) dla G implikuje wówczas istnienie punktu stałego dla tego działania. Z powyższego faktu i poprzedniego przykładu łatwo wywnioskować istnienie kolejnych przykładów grup nie posiadających własności (T): dowolna grupa posiadająca grupę liczb całkowitych jako iloraz nie będzie grupą sztywną. Tak więc dowolna grupa abelowa, czy ogólniej grupa posiadająca nieskończoną abelianizację, nie jest sztywna. W szczególności sztywne nie są też grupy wolne.

Jakie zatem istnieją przykłady grup sztywnych? Pierwszym przykładem są grupy skończone. Mając dane działanie grupy skończonej na przestrzeni Hilberta H możemy zdefiniować

$$\omega = 1/\#G \sum_{g\in G} \varphi(g)v$$

dla dowolnego $v \in H$. Innymi słowy ω jest środkiem ciężkości orbity v. Łatwo sprawdzić, że tak zdefiniowany w jest punktem stałym działania.

Powyższy przykład może służyć jako wyjaśnienie czy motywacja pojęcia sztywności, ponieważ istotą definicji własności (T) dla grup nieskończonych jest rodzaj zwartości, czy skończoności, zachowywany przez te grupy pomimo faktu, iż są one w rzeczywistości nieskończone.

Jakie są zatem przykłady nieskończonych grup sztywnych? Tych nie jest wiele, ze względu na to, że własność (T) jest własnością niezwykle silną. Pierwszymi przykładami były specjalne grupy liniowe wyższej rangi, czyli grupy

$$\mathrm{SL}_n(\mathbb{Z})$$

dla $n \geq 3$. Wskazane zostały one przez D. Kazhdana w 1966 w klasycznym już, 3-stronicowym artykule [1], z którego wywodzi się też samo pojęcie własności (T). Główną osią dowodu jest tu struktura algebraiczna grup liniowych, ich relacja z grupami Liego

$$\mathrm{SL}_n(\mathbb{R})$$

oraz znajomość teorii reprezentacji tych ostatnich.

Druga duża klasa przykładów nieskończonych grup sztywnych pojawiła się w drugiej połowie lat dziewięćdziesiątych ubiegłego wieku i w jej odnalezieniu duże zasługi mają polscy matematycy. Przykłady te to grupy automorfizmów pewnych kompleksów symplicjalnych, tak zwanych budynków typu

$$\tilde{A}_2.$$

Budynki te charakteryzują się dużą ilością rozgałęzień w wyższych wymiarach, jak w następującym przykładzie. Rozważmy płaszczyznę ze standardową triangulacją zadaną przez izometryczne trójkąty równoboczne – każda krawędź jest wówczas wspólna dla dwóch sympleksów. Teraz wyobraźmy sobie, że do każdej krawędzi przyczepiamy jeszcze jeden sympleks, który z kolei leży w nowej płaszczyźnie, triangulowanej identycznie jak poprzednia. Powtórzymy to dla każdej krawędzi, później każdej z krawędzi w „nowych" płaszczyznach i tak w nieskończoność. Kompleks symplicjalny, który otrzymamy w ten sposób jest właśnie przykładem budynku typu

$$\tilde{A}_2.$$

Pierwszy dowód własności (T) dla grup automorfizmów takich budynków podali Cartwright, Młotkowski i Steger [2] poprzez obliczenie widma operatora Laplace'a w maksymalnej C*-algebrze grupowej. Niedługo później w Ballmann i Świątkowski [3], Pansu [4] i Żuk [5] podali warunek konieczny na własność (T) w postaci warunku na widmo operatora Laplace'a pewnego skończonego grafu stowarzyszonego z grupą. Jeśli najmniejsza dodatnia wartość własna tego Laplasjanu jest ostro większa niż 1/2 wówczas grupa ma własność (T). Warunek ten można interpretować jako pewną formę dodatniej krzywizny dla grupy, ta krzywizna implikuje ograniczoność orbit i w efekcie znów środek ciężkości takiej orbity jest punktem stałym. Szerokie omówienie dowodów własności (T) dla obydwu powyższych klas przykładów można znaleźć w książce [10].

Najnowsza klasa przykładów została wskazana niedawno w pracach [6] i [7]. Są to grupy

$$\mathrm{Aut}(F_n)$$

automorfizmów grup wolnych na n generatorach dla $n \geq 5$. Pytanie czy grupy te mają własność (T) było otwarte od co najmniej lat dziewięćdziesiątych XX wieku. Metoda dowodu jest w tym przypadku całkiem nowa a pomysł na nią pochodzi z pracy [8] Narutaki Ozawy z 2016 roku. Ozawa pokazał, że własność (T) grupy G można scharakteryzować za pomocą warunku skończenie wymiarowego, a dokładniej w terminach istnienia macierzy dodatnio określonej spełniającej pewne równanie w pierścieniu grupowym $\mathbb{R}G$. W pracy [6] strategia ta została wykorzystana dla grupy

$$\mathrm{Aut}(F_5)$$

przy wsparciu metod optymalizacyjnych. Te ostatnie zostały wykorzystane do numerycznego znalezienia macierzy spełniającej odpowiednie równanie. Co prawda rozwiązanie numeryczne z natury nie jest dokładne, jednak w tym szczególnym przypadku dodatkowy argument wykorzystujący własności porządku na pierścieniu grupowym pozwala wywnioskować istnienie dokładnego rozwiązania, a co za tym idzie, własności (T). Najnowsze wyniki przedstawione w [7] pozwoliły zredukować problem własności (T) dla całej rodziny grup

$$\mathrm{Aut}(F_n),$$

gdzie

$$n \geq 6,$$

do jednego obliczenia numerycznego, znów w formie znalezienia dużej macierzy spełniającej zadane równanie w pierścieniu grupowym grupy

$$\mathrm{Aut}(\mathbb{F}_5).$$

Metody numeryczne rozwinięte w pracy [6] pozwoliły wykonać to obliczenie i uzyskać nową, nieskończoną rodzinę grup z własnością (T).

Wyniki te pozostawiają kilka otwartych pytań. Jednym z nich jest to czy

$$\mathrm{Aut}(F_4)$$

posiada własność (T) – jest to jedyny nierozstrzygnięty przypadek dla grup automorfizmów grup wolnych. Drugie pytanie to czy tzw. „mapping class grupy" posiadają własność (T). Mapping class grupa rozmaitości M to grupa homeomorfizmów $M \to M$, modulo relacja izotopii, czyli homotopii, w której każdy element jest homeomorfizmem. Czy metody obliczeniowe pozwolą rozwiązać te problemy?

LITERATURA

[1] Każdan, D. A. *On the connection of the dual space of a group with the structure of its closed subgroups*. Funkcional. Anal. i Priložen. 1 1967 71–74.

[2] D. I. Cartwright, W. Młotkowski, T. Steger, *Property (T) and $\tilde{A}_2$-groups*, Ann. Inst. Fourier (Grenoble) 44 (1994), no. 1, 213–248.

[3] W. Ballmann, J. Świątkowski, *On L^2-cohomology and property (T) for automorphism groups of polyhedral cell complexes*. Geom. Funct. Anal. 7 (1997), no. 4, 615–645.

[4] P. Pansu, *Sous-groupes discrets des groupes de Lie: rigidité, arithméticité*, Séminaire Bourbaki, Vol. 1993/94. Astérisque No. 227 (1995), Exp. No. 778, 3, 69–105.

[5] A. Żuk, *La propriété (T) de Kazhdan pour les groupes agissant sur les polyèdres*, C. R. Acad. Sci. Paris Sér. I Math. 323 (1996), no. 5, 453–458.

[6] M. Kaluba, P. W. Nowak, N. Ozawa, *Aut($\mathbb{F}_5$) has property (T)*, arXiv:1712.07167, Mathematische Annales, przyjęta.

[7] M. Kaluba, D. Kielak, P. W. Nowak, *On property (T) for Aut(F_n) and $SL_n(\mathbb{Z})$*, arXiv:1812.03456.

[8] N. Ozawa, *Noncommutative real algebraic geometry of Kazhdan's property (T)*, J. Inst. Math. Jussieu 15 (2016), no. 1, 85–90.

[9] K. Hartnett, *Best-Ever Algorithm Found for Huge Streams of Data*, Quanta Magazine, October 24, 2017, https://www.quantamagazine.org/best-ever-algorithm-found-for-huge-streams-of-data-20171024/

[10] B. Bekka, P. de la Harpe, A. Valette, *Kazhdan's property (T)*. New Mathematical Monographs, 11. Cambridge University Press, Cambridge, 2008. xiv+472.

{**Piotr W. Nowak** – polski matematyk, specjalizujący się w geometrii, topologii, geometrycznej teorii grup; nauczyciel akademicki. W 2003 ukończył studia magisterskie na Wydziale Matematyki, Informatyki i Mechaniki Uniwersytetu Warszawskiego. Za pracę magisterską otrzymał Nagrodę im. Marcinkiewicza 3 stopnia od Polskiego Towarzystwa Matematycznego. Stopień doktora uzyskał w 2008 na Vanderbilt University w Nashville w USA. Później, przez kolejne 4 lata pracował na Texas A&M University w College Station oraz w instytucie MSRI w Berkeley w Kalifornii oraz na Uniwersytecie w Oxfordzie w Wielkiej Brytanii. Pracując na uczelniach w Stanach Zjednoczonych kierował dwoma grantami National Science Foundation. Od 2011 pracuje w Instytucie Matematycznym Polskiej Akademii Nauk i na Uniwersytecie Warszawskim. W 2015 otrzymał prestiżowy grant European Research Council. Jest pierwszym polskim matematykiem, który otrzymał grant ERC. W 2016 roku obronił rozprawę habilitacyjną zatytułowaną *Kohomologie grup o współczynnikach w modułach Banacha*. Otrzymał za nią nagrodę Prezesa Rady Ministrów. Jest autorem ponad 30 artykułów. Wspólnie z Guoliangiem Yu wydał w 2012 roku książkę, będącą wprowadzeniem do geometrii dużej skali. Na międzynarodowych konferencjach i seminariach wygłosił ponad 150 zaproszonych wykładów na temat swoich wyników.}

MEDAL MŁODEGO UCZONEGO 2019 DLA PIOTRA ACHINGERA

Laudacja wygłoszona 25 czerwca 2019 r przez profesora Stanisława Janeczko z okazji uroczystego wręczenia Medalu za przełomowe rozwinięcie badań w zakresie topologii rozmaitości algebraicznych i ich zastosowań

Piotr Achinger urodził się 28 lipca 1986 roku w Warszawie. W roku 2005 ukończył warszawskie XIV Liceum Ogólnokształcące im. Stanisława Staszica. Już jako uczeń szkoły średniej zdobył pierwsze laury matematyczne. W roku 2004 dostał się do finału 55. Olimpiady Matematycznej, a rok później (2005) został laureatem II miejsca 56. OM i zdobywcą brązowego medalu na 46. Międzynarodowej Olimpiadzie Matematycznej. Uczestniczył także w 5. Zawodach Czesko-Polsko-Słowackich.

Jest absolwentem Wydziału Matematyki, Informatyki i Mechaniki UW, gdzie ukończył jednoczesne studia informatyczno-matematyczne. Uczestniczył w 13.,14.,15. i 16. Międzynarodowych Zawodach Matematycznych dla studentów, na których trzykrotnie zdobył pierwsze nagrody i raz drugą. W roku 2010 uzyskał tytuł magistra matematyki na podstawie pracy zatytułowanej „Frobenius push-forwards on quadrics". Promotorem był profesor Adrian Langer.

W tym samym roku przyznano mu stypendium doktoranckie w Hausdorff Center for Mathematics w Bonn w Niemczech. Po roku (w 2011) otrzymał International Fulbright Science&Technology Award, co umożliwiło mu kontynuację studiów doktoranckich w University of California at Berkeley, USA. W maju 2015 roku obronił pracę doktorską zatytułowaną „$K(\pi,1)$ spaces in algebraic geometry" przygotowaną pod kierunkiem profesora Arthura Edwarda Ogusa. Dwuletni staż podoktorski przyznany przez European Post Doctoral Institute spędził w Centrum Banacha w Warszawie oraz w Institut des Hautes Etudes Scientifiques w Paryżu (IHES, Bures-sur-Yvette). Od 2016 roku jest zatrudniony na stanowisku adiunkta w Zakładzie Algebry i Geometrii Algebraicznej Instytutu Matematycznego PAN w Warszawie.

Piotr Achinger zajmuje się topologią rozmaitości algebraicznych oraz geometrią algebraiczną (arytmetyczną) w dodatniej charakterystyce. Jego główne zainteresowania badawcze skupiają się wokół rozmaitości torycznych, kategorii pochodnych snopów koherentnych, rozmaitości abelowych, działań morfizmu Frobeniusa, spektralnej teorii grafów, algebry przemiennej i algebry homologicznej, w szczególności triangulowalnych kategorii osobliwości, modułów Cohena-Macaulaya, rozkładów macierzowych i teorii schematów. Wśród jego osiągnięć są między innymi istotne rozszerzenia rezultatów Faltingsa dotyczące gładkich schematów nad dyskretnym pierścieniem waluacyjnym oraz dowody szeregu twierdzeń dotyczących przestrzeni typu $K(\pi,1)$.

Piotr Achinger napisał o sobie: „Geometria algebraiczna – zajmuje się zbiorami rozwiązań układów równań wielomianowych. Ma bliskie związki z jednej strony z teorią liczb (np. dowód Wielkiego Twierdzenia Fermata), z drugiej strony – z fizyką (teoria strun, symetria lustrzana). Fakt, że współczesna geometria algebraiczna dostarcza jednorodnego języka, w którym można mówić o tak różnorodnych problemach, spowodował (stosunkowo niedawno) dość zaskakujący przepływ idei pomiędzy tymi pozornie niezwiązanymi dziedzinami. Różne aspekty mojej pracy wiążą się jakoś z umacnianiem tego pomostu".

Dodajmy, że od lat z pasją pomaga organizować kolejne edycje Olimpiady Matematycznej, współpracuje z Krajowym Funduszem na rzecz Dzieci, który przyznaje stypendia zdolnym młodym osobom z całej Polski i aktywnie uczestniczy w prowadzeniu Wakacyjnych Warsztatów Wielodyscyplinarnych, coroczne wydarzenie organizowane przez studentów UW pod patronatem Koła Pasjonatów Matematyki UW i Studenckiego Koła Fizyki UW, przeznaczone dla licealistów zainteresowanych matematyką, informatyką, fizyką lub astronomią. „Tego typu wsparcie ma w moim odczuciu większy wpływ na rozwój młodych talentów niż konkursy matematyczne. I pozwala na naukę bez określonego celu. Uważam, że to jest bardzo ważne. Czasem pozorna bezcelowość może przynieść najwięcej dobrego" – uważa Piotr Achinger. Jego postawa wokół młodych adeptów nauki w szczególności matematyki niezwykle harmonizuje ze spostrzeżeniem Einsteina: „W rzeczywistości nie odbiega daleko od cudu to, że nowoczesne metody nauczania jeszcze zupełnie nie zdławiły świętej ciekawości dociekań; bowiem ta delikatna roślinka poza potrzebą bodźca wymaga głównie wolności, bez której niewątpliwie stanie się wrakiem i ruiną".

Piotr Achinger w 2016 roku został laureatem Nagrody im. Kazimierza Kuratowskiego za prace z geometrii algebraicznej w dodatniej charakterystyce. Nagroda ustanowiona w 1981 roku przez Zofię Kuratowską, Instytut Matematyczny PAN i Polskie Towarzystwo Matematyczne przyznawana jest naukowcom, którzy nie ukończyli 30 lat do końca roku poprzedzającego przyznanie nagrody i którzy nie są laureatami nagród PTM (z wyłączeniem nagród PTM dla młodych matematyków), ani też Nagrody Naukowej Wydziału III PAN.

Kolejnym osiągnięciem tego niezwykle zdolnego matematyka było uzyskanie w 2018 roku ERC Starting Grant. Prestiżowy grant w wysokości 1 007 500 euro jest drugim grantem ERC realizowanym w Instytucie Matematycznym Polskiej Akademii Nauk i jednocześnie drugim matematycznym grantem ERC w Polsce. Dzięki tym funduszom matematyk realizuje projekt dotyczący topologii rozmaitości algebraicznych i ich zastosowań w arytmetyce, fizyce matematycznej i równaniach różniczkowych.

Piotr Achinger jest wybitnym matematykiem o niezwykłej zdolności twórczego poruszania się po najtrudniejszych i najpiękniejszych obszarach współczesnej matematyki. Jest wybitną indywidualnością wyróżniającą się samodzielnością koncepcji twórczych i ukazującą w całej rozciągłości, swoimi osiągnięciami, głęboką prawdę sentencji „Qantum scimus gutta est ignoramus mare", wyrytej na rewersie Medalu Młodego Uczonego.

Profesor Stanisław Janeczko
Durektor CSZ PW

Typy homotopii w geometrii algebraicznej

Piotr Achinger

Typ homotopii przestrzeni topologicznej opisuje własności jej „kształtu", zapominając geometryczne informacje nieistotne z punktu widzenia topologii algebraicznej (np. litery A i O mają ten sam typ homotopii). Istotne niezmienniki, takie jak grupa podstawowa lub grupy kohomologii, zależą wyłącznie od typu homotopii przestrzeni.

Geometria algebraiczna zajmuje się rozmaitościami algebraicznymi oraz innymi obiektami geometrycznymi zdefiniowanymi w algebraiczny sposób. Metody topologii algebraicznej okazują się być bardzo pomocne w ich badaniu, ponieważ odpowiednie niezmienniki (grupy kohomologii, grupa podstawowa) mają dużo więcej struktury niż w topologii.

Niniejsze notatki stanowią zaadaptowaną na potrzeby Biuletynu CSZ PW Profundere Scientiam wersję wygłoszonego przeze mnie XXVI Wykładu im. Wojtka Pulikowskiego na Uniwersytecie Adama Mickiewicza w Poznaniu 6 czerwca 2019 r. Uzupełniają one wykład o matematyczne szczegóły (definicje, precyzyjne sformułowania twierdzeń) oraz wyniki, których zrozumienie wymaga nieco więcej znajomości matematyki wyższej. Te trudniejsze fragmenty oznaczono gwiazdką (*).

Zgodnie z powszechnym poglądem, *geometria* i *algebra* wzajemnie się uzupełniają. Algebra jest zazwyczaj lepiej dostosowana do *ścisłych rozumowań* niż geometria (jak potwierdzi zapewne każdy kto kiedykolwiek sprawdzał zadania na Olimpiadzie Matematycznej czy podobnych zawodach). Z kolei geometria jest często bliższa ludzkiej intuicji.

Topologia algebraiczna może być widziana jako algebraizacja geometrii. Dany obiekt geometryczny – przestrzeń topologiczna – zamienia się na obiekt algebraiczny (np. grupę podstawową czy grupy homologii). Zazwyczaj nie sposób z tak otrzymanego niezmiennika odzyskać badany obiekt – typowe niezmienniki nie odróżniają od siebie homotopijnie równoważnych przestrzeni. *Typ homotopii* przestrzeni to jej klasa abstrakcji względem relacji homotopijnej równoważności.

Analogicznie, *geometria algebraiczna* to geometryzacja algebry (dokładniej, teorii pierścieni przemiennych). Jednak, w odróżnieniu od topologii algebraicznej, w geometrii algebraicznej geometria i algebra stają się tym samym. Pozwala to naprzemiennie stosować techniki z obu dziedzin.

Poniżej opisze, w jaki sposób metody topologii algebraicznej znajdują zastosowanie – pośrednio i bezpośrednio – w geometrii algebraicznej. Podstawowym pytaniem jest: *jak z obiektem geometrycznym zdefiniowanym w algebraiczny sposób, jakim jest rozmaitość algebraiczna, stowarzyszyć zdefiniowany w algebraiczny sposób typ homotopii?* Na to pytanie nie odpowiem, ale postaram się wytłumaczyć, dlaczego jest ono istotne.

1. Co to jest rozmaitość algebraiczna?

Niech K będzie ciałem. Dla ustalenia uwagi, wymieńmy kilka ciał, które pojawiają się najczęściej:

- → $\mathbf{C}$, ciało liczb zespolonych,
- → $\mathbf{R}$, liczby rzeczywiste,
- → $\mathbf{Q}$, liczby wymierne, ciała skończone: $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$, $\mathbf{F}_{p^n}$,
- → $\mathbf{Q}_p$, liczby p-adyczne,
- → $\mathbf{C}(t)$, funkcje wymierne zmiennej zespolonej,
- → algebraiczne domknięcia powyższych: $\overline{\mathbf{Q}}$, $\overline{\mathbf{F}}_p, \ldots$

Definicja. *Afiniczną rozmaitością algebraiczną* nad K nazywamy zbiór X rozwiązań $(x_1, \ldots, x_n)$ układu równań wielomianowych

$$\begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ \ldots \\ f_r(x_1, \ldots, x_n) = 0 \end{cases}$$

o współczynnikach w K.

Dwie natychmiastowe uwagi do powyższej definicji. Po pierwsze, ściśle rzecz biorąc, powyżej zdefiniowaliśmy „schemat afiniczny skończonego typu nad K" (w tym tekście nie będziemy rozwodzili się nad różnicą pomiędzy rozmaitościami i schematami). Po drugie, *celowo* nie wspomnieliśmy, do jakiego zbioru mają należeć wartości zmiennych $x_1, \ldots, x_n$! W istocie, interesują nas rozwiązania $(x_1, \ldots, x_n) \in L^n$ dla każdego rozszerzenia L ciała K, a nawet $(x_1, \ldots, x_n) \in R^n$ dla każdej K-algebry R. (Gdybyśmy rozważali tylko rozwiązania w K, wtedy np. dla $K = \mathbf{R}$ bylibyśmy zmuszeni uznać rozmaitość zadaną równaniem

$$x_1^2 + x_2^2 = -1$$

za zbiór pusty, podczas gdy równanie to ma rozwiązania nad $L = \mathbf{C}$! Dla ciał algebraicznie domkniętych ten problem jest mniej istotny.) Wprowadźmy następujące oznaczenie na zbiór rozwiązań o wartościach w R:

$$X(R) = \{(x_1, \ldots, x_n) \in R^n \mid f_i(x_1, \ldots, x_r) = 0 \text{ dla } i = 1, \ldots, r\}.$$

Ewaluacja f_i w $(x_1, \ldots, x_n)$ ma sens ponieważ współczynniki wielomianów f_i leżą w K a R zawiera K. Zauważmy też, że jeżeli $f: R \to R'$ jest homomorfizmem K-algebr, to rozwiązania o wartościach w R dają rozwiązania o wartościach w R', innymi słowy mamy indukowane odwzorowanie

$$f: X(R) \to X(R').$$

Musimy pamiętać nie tylko zbiory $X(R)$ dla wszystkich R, ale też powyższe odwzorowania dla wszystkich f. Innymi słowy, przez X rozumiemy odpowiedni *funktor* z kategorii K-algebr do kategorii zbiorów. To ułatwia zdefiniowanie morfizmu rozmaitości – jest to po prostu transformacja naturalna funktorów.

Przykład (Przestrzenie afiniczne). Najprostsze przykłady rozmaitości afinicznych to *przestrzenie afiniczne*

$$X = \mathbf{A}^n_K, \quad X(R) = R^n.$$

(Mamy tu n zmiennych oraz 0 równań.) Z definicji, każda afiniczna rozmaitość algebraiczna jest (domknięta) podrozmaitością przestrzeni afinicznej.

Przykład (Okrąg jednostkowy). Przykład znany ze szkoły: okrąg jednostkowy X na płaszczyźnie ($\mathbf{A}^2_{\mathbf{R}}$) dany jest równaniem

$$x^2 + y^2 = 1.$$

(W przypadku dwu zmiennych będziemy je nazywali x i y zamiast x_1 i x_2)

Przykład (Krzywe Fermata). *Krzywą Fermata* (afiniczną) nazywamy rozmaitość X_n ($n \geq 2$) zadaną równaniem

$$x^n + y^n = 1.$$

Jak wspomnieliśmy, w geometrii algebraicznej kluczowe jest rozpatrywanie rozwiązań nad różnymi ciałami. Zobaczmy tę różnorodność teraz w praktyce na powyższych przykładach:

- $X_2(\mathbf{R})$ to okrąg jednostkowy,
- $X_2(\mathbf{Q})$ odpowiada trójkom Pitagorejskim (np. $(3/5, 4/5)$),
- $X_n(\mathbf{Q})$ jest opisane przez Wielkie Twierdzenie Fermata,
- $X_n(\mathbf{C})$ to (nakłuta) powierzchnia Riemanna genusu $g = (n-1)(n-2)/2$,
- $X_3(\mathbf{F}_p)$ to skończona grupa abelowa mająca (być może) zastosowanie w kryptografii,
- liczby $|X_n(\mathbf{F}_{p^n})|$ dla wszystkich p^n, spakowane wygodnie w tzw. L-funkcje, hipotetycznie zawierają bardzo głębokie informacje teorioliczbowe (np. hipoteza Bircha i Swinnertona-Dyera).

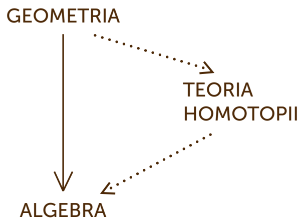
Ogólne (nieafiniczne) *rozmaitości algebraiczne* (czy „schematy skończonego typu") są zbudowane z rozmaitości afinicznych, podobnie jak rozmaitości topologiczne i różniczkowe są sklejone z otwartych podzbiorów w $\mathbf{R}^n$. Najistotniejsza rodzina przykładów rozmaitości nieafinicznych to *rozmaitości rzutowe*, zadane za pomocą układów równań wielomianowych *jednorodnych* w przestrzeni rzutowej. Dla przykładu, rzutowa krzywa Fermata:

$$\overline{X}_n = \{X^n + Y^n = Z^n\} = X_n \cup \{\infty\} \subseteq \mathbf{P}^2_K.$$

2. Topologia algebraiczna i geometria algebraiczna

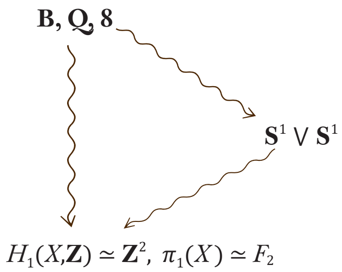
Opiszemy teraz jak i dlaczego metody topologii algebraicznej są przydatne w badaniu rozmaitości algebraicznych.

Zacznijmy od krótkiego zarysu działania topologii algebraicznej. Obiektowi geometrycznemu (sensownej przestrzeni topologicznej) przyporządkowuje się obiekt algebraiczny (np. grupę podstawową czy grupy kohomologii). Przestrzeniom homotopijnie równoważnym przypisuje się ten sam obiekt; utożsamiając ze sobą przestrzenie homotopijnie równoważne (oraz homotopijne przekształcenia)



zamieniamy przestrzenie na ich odpowiednie typy homotopii.

Przykład. Symbole **B**, **Q**, **8** posiadają ten sam typ homotopii, często opisywany jako „bukiet dwu okręgów" ($\mathbf{S}^1 \vee \mathbf{S}^1$). Ich ciekawe niezmienniki to pierwsza grupa homologii $H_1(\mathbf{S}^1 \vee \mathbf{S}^1, \mathbf{Z}) \simeq \mathbf{Z}_2$ oraz grupa podstawowa $\pi_1(\mathbf{S}^1 \vee \mathbf{S}^1) \simeq F_2$ (grupa wolna na dwu generatorach).



2.1. Bezpośrednie zastosowanie topologii algebraicznej w geometrii algebraicznej

Aby móc zastosować maszynerię topologii algebraicznej do badania rozmaitości algebraicznej X nad K, musimy wyprodukować sensowną przestrzeń topologiczną odpowiadającą X. W geometrii algebraicznej często bada się odpowiadającą X przestrzeń topologiczną $X(\overline{K})$ wyposażoną w tzw. topologie Zariskiego (zbiory domknięte w topologii Zariskiego to te zadane przez równania wielomianowe), jednak jest ona zbyt patologiczna z punktu widzenia topologii algebraicznej.

Najprostsze rozwiązanie tego problemu jest możliwe, jeżeli ciało K zanurza się w $\mathbf{C}$, bo wtedy zbiór $X(\mathbf{C})$, z topologią indukowaną ze zwykłej topologii $\mathbf{C}^n$, jest sensowną przestrzenią topologiczną:

$$\iota : K \hookrightarrow \mathbf{C} \rightsquigarrow X(\mathbf{C}) \rightsquigarrow H_n(X(\mathbf{C}), \mathbf{Z}, \pi_1(X(\mathbf{C})), \ldots$$

Jest kilka problemów z powyższym podejściem:

- Zanurzenie ι może nie istnieć, np. jeśli ciało K ma dodatnią charakterystykę.
- Naturalny sposób obejścia powyższego np. dla $K = \mathbf{F}_p$ to znalezienie „podniesienia" $\hat{X}$ rozmaitości X do charakterystyki zero i rozważenie $\hat{X}(\mathbf{C})$. Jak pokazał Serre [Ser61], takie podniesienia nie zawsze istnieją.
- Jak również pokazał Serre [Ser64], dla ustalonego X, typ homotopii $X(\mathbf{C})$ może zależeć od wyboru ι.
- W związku z powyższym, nie istnieje naturalne działanie grupy Galois $G_K = \mathrm{Gal}(\overline{K}/K)$ na $X(\mathbf{C})$, jego typie homotopii, ani na jego grupach (ko)homologii.

2.2. Topologia étale i kohomologie étale (Artin–Grothendieck).

W związku z powyższymi problemami, naturalne wydaje się następujące podejście. Zamiast stosować metody topologii algebraicznej bezpośrednio do pewnej przestrzeni otrzymanej z X, spróbujmy stworzyć w algebraiczny sposób narzędzia analogiczne do narzędzi topologii algebraicznej. Najbardziej skuteczną próbą tego typu było skonstruowanie przez Artina i Grothendiecka tzw. topologii étale (jest to tzw. topologia Grothendiecka, czyli uogólnienie pojęcia przestrzeni topologicznej istotne z punktu widzenia teorii snopów) oraz kohomologii étale [SGA73]. Nie będziemy tutaj opisywać ich konstrukcji, ale podamy kilka własności.

Jednym z mankamentów na wstępie tej teorii jest konieczność wyboru pomocniczej liczby pierwszej ℓ (jeżeli ciało K ma charakterystykę $p > 0$, wymagamy $\ell \neq p$). Jeżeli X jest rozmaitością algebraiczną nad K, wtedy kohomologie étale X to przestrzenie liniowe nad ciałem liczb ℓ-adycznych $\mathbf{Q}_\ell$:

$$X \rightsquigarrow H^n(X, \mathbf{Q}_\ell)$$

(mówiąc ściślej, rozważamy grupy $H^n_{\mathrm{et}}(X_{\overline{K}}, \mathbf{Q}_\ell)$). Są one wyposażone w naturalne ciągłe działanie grupy Galois $G_K = \mathrm{Gal}(\overline{K}/K)$.

Aby przekonać się, że grupy kohomologii étale są sensownym pomysłem, należy je porównać z grupami kohomologii $X(\mathbf{C})$ w przypadku $K = \mathbf{C}$. Istotnie, Artin udowodnił, że istnieje naturalny izomorfizm

$$H^n(X, \mathbf{Q}_\ell) \simeq H^n(X(\mathbf{C}), \mathbf{Q}_\ell),$$

gdzie po lewej stronie mamy algebraicznie zdefiniowane grupy kohomologii étale, a po prawej – grupy kohomologii (singularnych) przestrzeni topologicznej $X(\mathbf{C})$ zdefiniowane klasycznie przez topologów algebraicznych kilkadziesiąt lat wcześniej.

Jakie są zastosowania grup kohomologii étale? Chyba najważniejszym elementem ich struktury jest działanie grupy Galois, które ma tendencje do

„pamiętania" informacji o punktach wymiernych $X(K)$ oraz podrozmaitościach algebraicznych (jest to sformułowane precyzyjnie w hipotezach Weila – o których opowiemy za chwile – oraz Tate'a).

Zakończę ten temat następującą ogólną uwagą. Metody topologii algebraicznej stają sie *jeszcze potężniejsze* w badaniu rozmaitości algebraicznych (w odróżnieniu od przestrzeni topologicznych), ponieważ odpowiednie niezmienniki (np. grupy kohomologii) są wyposażone w dużo więcej struktury (np. wspomniane wyżej działanie grupy Galois. Jest to prawda nawet dla $K = \mathbf{C}$, gdzie grupy kohomologii są wyposażone w tzw. *struktury Hodge'a*. Zgodnie z jednym z Problemów Milenijnych – hipoteza Hodge'a – struktury te pamiętają, które klasy w kohomologiach pochodzą od podrozmaitości algebraicznych (ściślej – „wymiernych cykli algebraicznych").

2.3. Hipotezy Weila

Jedną z głównych motywacji dla skonstruowania kohomologii étale przez Artina i Grothendiecka były hipotezy postawione kilkanaście lat wcześniej przez Weila. Dotyczą one liczenia punktów na rozmaitościach algebraicznych nad ciałami skończonymi.

Niech zatem $K = \mathbf{F}_q$ (gdzie $q = p^e$ dla pewnej liczby pierwszej p oraz $e \geq 1$) będzie ciałem skończonym i niech X będzie rozmaitością algebraiczną (((*)) dokładniej: rozmaitością algebraiczną gładką i rzutową). Zastanawiamy się, ile punktów leży na X nad ciałem K i jego rozszerzeniami $L = \mathbf{F}_{q^m}$:

$N_m := |X(\mathbf{F}_{q^m})| = ?$

O dziwo, to pytanie ma silny związek z kohomologiami! Liczby N_m jest wygodnie upakować w funkcje tworzącą, tzw. *funkcję dzeta Hassego-Weila*, następującej postaci

$$Z(X,t) = \exp\left(\sum_{m\geqslant 1} \frac{|X(\mathbf{F}_{q^m})|}{m} t^m\right)$$

Przykład. Dla płaszczyzny rzutowej $X = \mathbf{P}^2_K$ (którą można rozłożyć na $\mathbf{A}^0_K \cup \mathbf{A}^1_K \cup \mathbf{A}^2_K$, liczba punktów nad ciałem $\mathbf{F}_{q^m}$ wynosi

$$N_m = 1 + q + q^2.$$

Zatem

$$Z(X,t) = \frac{1}{(1-t)(1-qt)(1-q^2t)}.$$

Zauważmy jednocześnie, że wymiary grup kohomologii $H^{2i}(\mathrm{C}P^2,\mathbf{Q})$ wynoszą kolejno 1, 1, 1 (nieparzyste grupy kohomologii znikają). Zgodnie z hipotezami Weila, liczby N_m kodują wymiary grup kohomologii.

Przykład. Niech $X = X_n$ będzie (rzutowa) krzywa Fermata (zakładamy tutaj, że p nie dzieli n, w przeciwnym przypadku ta krzywa nie jest gładka). Wtedy

$$Z(X,t) = \frac{P(t)}{(1-t)(1-qt)},$$

gdzie $P(t)$ jest wielomianem unormowanym o współczynnikach całkowitych stopnia $2g = \dim H^1(X(\mathbf{C}),\mathbf{Q})$, którego wszystkie pierwiastki zespolone t spełniają

$$|t| = \sqrt{q}.$$

Związek liczb N_m a więc i funkcji dzeta z kohomologiami jest zawarty *we wzorze śladu Lefschetza– Grothendiecka*:

$$N_m = \sum_{i\geqslant 0} (-1)^i \operatorname{Tr}(F^m \mid H^m(X, \mathbf{Q}_\ell)).$$

Tutaj F jest odwzorowaniem Frobeniusa ($F(a) = a^{1/q}$), które jest generatorem grupy Galois ciała K. Zauważmy, że grupa Galois $\mathrm{G}_{\mathbf{F}_q} = \langle \mathrm{F} \rangle$ działa na $X(\overline{\mathbf{F}}_q)$ i punkty stałe to dokładnie $X(\mathbf{F}_q)$, podobnie punkty stałe działania podgrupy $\mathrm{G}_{\mathbf{F}_{q^m}} = \langle F_m \rangle$ to $X(\mathbf{F}_{q^m})$.

Możemy teraz sformułować hipotezy Weila, które uogólniają powyższe przykłady.

Twierdzenie (Dwork, Grothendieck, Deligne). *Dla dowolnej rozmaitości algebraicznej gładkiej i rzutowej X nad ciałem skończonym $K = \mathbf{F}_q$, funkcja dzeta $Z(X, t)$ ma postać*

$$Z(X,t) = \prod_{i=0}^{2\dim X} P_i(t)^{(-1)^{i+1}}$$

gdzie P_i są wielomianami unormowanymi o współczynnikach całkowitych stopnia $\deg P_i = \dim H^i(X,\mathbf{Q}_\ell)$. *Ich pierwiastki zespolone*

α_{ij}, $0 \leq i \leq \dim X$, $1 \leq j \leq \dim H^i(X, \mathbf{Q}_\ell)$ *spełniają*

$|\alpha_{ij}| = q^{i/2}$. („*hipoteza Riemanna*" [Del74])

W szczególności, liczby N_m można wyznaczyć za pomocą pierwiastków α_{ij}:

$$N_m = \sum_{i=0}^{\dim X} (-1)^i \sum_{j=1}^{\deg P_i} \alpha_{ij}^m.$$

A posteriori, znając wszystkie N_m, znamy liczby α_{ij} (znane ćwiczenie), a zatem znamy wymiary grup kohomologii. W szczególności, licząc punkty nad ciałami skończonymi jesteśmy w stanie powiedzieć coś o topologii rozmaitości zespolonych, co jest dość zaskakujące! Na odwrót, znajomość topologii daje nam oszacowanie na liczbę składników w powyższej sumie, a zatem na N_m.

3. Typ homotopii étale

Po zdefiniowaniu i zbadaniu kohomologii étale (oraz związanej z nimi *grupy podstawowej étale* $\pi_1^{\mathrm{et}}(X)$) powstało następujące pytanie: czy istnieje, dla danej rozmaitości algebraicznej X, naturalna „przestrzeń" (czy: typ homotopii w klasycznym znaczeniu tego słowa), której grupy kohomologii są izomorficzne z grupami kohomologii étale, i której grupa podstawowa zgadza się z grupą podstawową étale?

Na to pytanie wkrótce odpowiedzieli Artin i Mazur [AM69], którzy skonstruowali *typ homotopii étale* $\Pi^{\mathrm{et}}_\infty(X)$:

$$X \rightsquigarrow \Pi^{\mathrm{et}}_\infty(X) \rightsquigarrow H^n(X, \mathbf{Q}_\ell), \pi_1^{\mathrm{et}}(X), \ldots$$

(Formalnie, jest to pro-obiekt (pewna granica odwrotna) w kategorii typów homotopii.)

Aby przybliżyć czytelnikowi nieco idee konstrukcji typu homotopii étale, rozważmy wpierw, w jaki sposób możemy zrozumieć typ homotopii rozmaitości topologicznej M. Z definicji, M można pokryć ściągalnymi podzbiorami otwartymi:

$$M = \bigcup_{i\in I} U_i, \quad U_i \overset{\text{homeo}}{\simeq} \mathbf{R}^n.$$

Ich przecięcia $U_i \cap U_j$ niekoniecznie są ściągalne, ale można je z kolei pokryć ściągalnymi podzbiorami otwartymi:

$$U_i \cap U_j = \bigcup_{k\in I_{i,j}} U_{ijk}, \quad U_{ijk} \overset{\text{homeo}}{\simeq} \mathbf{R}^n.$$

Jeżeli będziemy odpowiednio skrupulatni jeśli chodzi o notacje, zbiory indeksów

$$I_0 = I, \quad I_1 = \coprod_{i,j\in I} I_{i,j}, \quad \ldots$$

wraz z naturalnymi odwzorowaniami pomiędzy nimi (np. istnieją dwa naturalne odwzorowania $I_1 \to I_0$, posyłające element $k \in I_{ij}$ na i lub na j), będą stanowiły obiekt kombinatoryczny zwany *zbiorem symplicjalnym* (((*)) formalnie, zbiór symplicjalny to funktor kontrawariantny z kategorii niepustych skończonych porządków liniowych do kategorii zbiorów):

$$I_\bullet \quad : \quad I_0 \leftleftarrows I_1 \;\substack{\longleftarrow\\\longleftarrow\\\longleftarrow}\; I_2 \quad \cdots$$

Zbiór symplicjalny $I_\bullet$ pamięta typ homotopii rozmaitości M.

Przykład. Okrąg jednostkowy ($\mathbf{S}^1 \subseteq \mathbf{C}$ można pokryć dwoma ściągalnymi zbiorami otwartymi

$$\mathbf{S}^1 = U_+ \cup U_-, \quad U_+ = \{\mathrm{Re}\, z > -1\}, \quad U_- = \{\mathrm{Re}\, z < 1\}.$$

Ich przecięcie $U_+ \cap U_-$ składa się z dwu ściągalnych składowych. Otrzymany zbiór symplicjalny ma po dwa niezdegenerowane sympleksy wymiaru 0 i 1.

(*) W kontekście rozmaitości algebraicznych, powyższe podejście napotyka na problem: nie są one „lokalnie ściągalne" w żaden naturalny sposób. Aby to przeskoczyć, Artin i Mazur rozpatrują wszystkie możliwe „hipernakrycia" (tj. systemy pokryć X „zbiorami otwartymi étale", wraz z pokryciami przecięć par tych zbiorów, i tak dalej, bez założenia ściągalności). Jeżeli

$$U_\bullet \quad : \quad U_0 \;\substack{\longleftarrow\\ \longrightarrow\\ \longleftarrow}\; U_1 \;\substack{\longleftarrow\\ \longrightarrow\\ \longleftarrow\\ \longrightarrow\\ \longleftarrow}\; U_2 \quad \cdots$$

jest takim hipernakryciem, zamieniamy każdą rozmaitość U_n zbiorem jej spójnych składowych $I_n = \pi_0(U_n)$, otrzymując zbiór symplicjalny $I_\bullet$, który oczywiście zależy od wyboru $U_\bullet$, nawet z dokładnością do homotopii. Idea polega na rozpatrywaniu wszystkich takich zbiorów symplicjalnych $I_\bullet$, indeksowanych zbiorem wszystkich hipernakryć $U_\bullet$ rozmaitości X. To daje szukany pro-obiekt $\Pi_\infty^{\mathrm{et}}(X)$ w kategorii homotopii (Artin–Mazur), który przy małej modyfikacji pochodzi od pro-obiektu w kategorii zbiorów symplicjalnych (Friedlander [Fri82]).

(*) W przypadku $K = \mathbf{C}$, mamy naturalny morfizm typów homotopii

$$X(\mathbf{C}) \to \Pi_\infty^{\mathrm{et}}(X),$$

który indukuje izomorfizm na „uzupełnieniach proskończonych" obu przestrzeni, uogólniając twierdzenie porównawcze dla kohomologii.

4. Charakterystyka 0 i charakterystyka p

Jak widzieliśmy, geometria algebraiczna nad ciałami charakterystyki zero ma tę zaletę, że (zazwyczaj) można założyć, że dane ciało to ciało $\mathbf{C}$ liczb zespolonych, co pozwala nam na korzystanie bezpośrednio z metod topologii algebraicznej, analizy, geometrii symplektycznej itd. Podobnie, geometria algebraiczna nad ciałami charakterystyki dodatniej pozwala na korzystanie z właściwych sobie metod: morfizmu Frobeniusa $x \mapsto x^p$ oraz liczenia punktów nad ciałami skończonymi.

Wiele głębokich wyników w geometrii algebraicznej nie posiada znanych dowodów nie uciekających się do jednej z wymienionych metod, i jest pożądane, żeby w danym kontekście można było użyć kilku z nich. Służą do tego dwie metody „zmiany charakterystyki": *redukcja modulo p* oraz *podnoszenie do charakterystyki zero*.

Jakkolwiek redukcję modulo p zawsze można wykonać z powodzeniem, jeżeli tylko p wziąć odpowiednio duże, tak (jak już wspomnieliśmy) podnoszenie do charakterystyki zero nie zawsze jest możliwe. Jest to związane z pewnymi patologiami natury topologicznej na rozmaitościach algebraicznych w dodatniej charakterystyce.

Typy homotopii rozmaitości nad ciałami charakterystyki dodatniej często mają bardzo dziwne własności. Dla przykładu:

- Prosta afiniczna $\mathbf{A}^1_K$ (K algebraicznie domknięte charakterystyki p) nie jest jednospójna (ma nietrywialne nakrycia).
  Najprostszym przykładem jest *nakrycie Artina–Schreiera*:
  $$f(t) = t - t^p \quad : \mathbf{A}^1_K \to \mathbf{A}^1_K.$$
  Aby sprawdzić, że to jest nakrycie, liczymy pochodną:
  $$f'(t) = 1 - pt^{p-1} = 1.$$
  Zatem prosta afiniczna nakrywa sama siebie p-krotnie! (*) Zjawisko jest to związane z tzw. dzikim rozgałęzieniem (ramifikacją) morfizmu f w nieskończoności.
- Jak pokazał Raynaud [Ray94], grupa podstawowa $\pi_1(\mathbf{A}^1_K)$ jest bardzo duża: każda grupa skończona G która nie posiada nietrywialnych ilorazów stopnia względnie pierwszego z p (na przykład, prawie wszystkie nieabelowe grupy proste) jest ilorazem $\pi_1(\mathbf{A}^1_K)$!
- Jak pokazali Holschbach–Schmidt–Stix [HSS14], nie istnieją „ściągalne" rozmaitości nad ciałem charakterystyki dodatniej inne niż punkt.

Następujące ogólne twierdzenie jest nieco zaskakujące – mówi ono, że grupa podstawowa jest w pewnym sensie „odpowiedzialna" za wszystkie te komplikacje.

Twierdzenie (A. 2017 [Ach17]). *Typ homotopii étale dowolnej rozmaitości algebraicznej nad ciałem charakterystyki dodatniej jest przestrzenią* $K(\pi,1)$.

Z definicji, przestrzeń $K(\pi,1)$ to taka przestrzeń Y, której wszystkie wyższe grupy homotopii $\pi_n(Y)$ $(n \geq 2)$ są zerowe. Typ homotopii takiej przestrzeni jest jednoznacznie wyznaczony przez jej grupę podstawową $\pi_1(Y)$.

Przykładowym wnioskiem z tego twierdzenia jest fakt, ze grupy podstawowe étale $\pi_1(\mathbf{A}^n_K)$ oraz $\pi_1(\mathbf{A}^m_K)$ nie są izomorficzne dla $n \neq m$.

(*) Problem otwarty: Niech K będzie ciałem doskonałym charakterystyki p (tj. dowolny jego element jest p-tą potęgą). Czy grupa $\pi_1(\mathbf{A}^1_K)$ wyznacza ciało K z dokładnością do izomorfizmu?

LITERATURA

[Ach17] Piotr Achinger, *Wild ramification and K(π,1) spaces*, Invent. Math. 210 (2017), no. 2, 453–499. MR 3714509

[AM69] M. Artin and B. Mazur, *Étale homotopy, Lecture Notes in Mathematics*, No. 100, Springer-Verlag, Berlin-New York, 1969. MR 0245577 (39 #6883)

[Del74] Pierre Deligne, *La conjecture de Weil*. I, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 273–307. MR 0340258

[Fri82] Eric M. Friedlander, *Étale homotopy of simplicial schemes*, Annals of Mathematics Studies, vol. 104, Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1982. MR 676809

[HSS14] Armin Holschbach, Johannes Schmidt, and Jakob Stix, *Étale contractible varieties in positive characteristic*, Algebra Number Theory 8 (2014), no. 4, 1037–1044. MR 3248993

[Ray94] M. Raynaud, *Revêtements de la droite affine en caractéristique p > 0 et conjecture d'Abhyankar*, Invent. Math. 116 (1994), no. 1-3, 425–462. MR 1253200

[Ser61] Jean-Pierre Serre, *Exemples de variétés projectives en caract éristique p non relevables en caractéristique zéro,* Proc. Nat. Acad. Sci. U.S.A. 47 (1961), 108–109. MR 0132067

[Ser64] *Éxemples de variétés projectives conjuguées non homéomorphes*, C. R. Acad. Sci. Paris 258 (1964), 4194–4196. MR 0166197

[SGA73] *Théorie des topos et cohomologie étale des schémas*. Tome 3, Lecture Notes in Mathematics, Vol. 305, Springer-Verlag, Berlin- New York, 1973, S´eminaire de G´eom´etrie Alg´ebrique du Bois- Marie 1963–1964 (SGA 4), Dirig´e par M. Artin, A. Grothendieck et J. L. Verdier. Avec la collaboration de P. Deligne et B. Saint-Donat MR 0354654 (50 #7132)

{**Piotr Achinger** - Instytut Matematyczny Polskiej Akademii Nauk (PAN) Zakład Algebry i Geometrii Algebraicznej. Laureat ERC Starting Grant 2018 - drugiego matematycznego grantu ERC realizowanego w Polsce. W 2019 r. nagrodzony Medalem Młodego Uczonego Politechniki Warszawskiej za przełomowe rozwinięcie badań w zakresie topologii rozmaitości algebraicznych i ich zastosowań}

jego trwania niemal cała aktywność intelektualna naszej populacji orbitowała wokół kwestii przydatności do spożycia okolicznej fauny i flory. Być może z dumą myślimy o wytworach naszej cywilizacji. Warto jednak pamiętać, że wszystko, co człowiek kiedykolwiek zbudował, z piramidami, wieżą Eiffla oraz Jezusem ze Świebodzina włącznie, zmieściłoby się wewnątrz kostki sześciennej pięć na pięć na pięć kilometrów. Co prawda niektórzy twierdzą, *że wybudowali sobie pomnik trwalszy niż ze spiżu, strzelający nad ogrom królewskich piramid*. Ja tam jednak nie oburzam się na słynnego antropologa, Desmonda Morrisa, który nasz gatunek określał słodkim mianem *nagiej małpy*.

A jednak całkiem niedawno naga małpa doznała gwałtownej erekcji intelektualnej. W ciągu zaledwie czterystu lat od nasilenia się jej objawów naga małpa zdołała wystrzelić swojego przedstawiciela na Księżyc, zbudować komputer oraz dokonać teleportacji kwantowej stanu pojedynczego fotonu na odległość ponad stu kilometrów. A w wolnej chwili zbadała dziedziczność tego, czy krowa, spożywając trawę, kręci swoją żuchwą zgodnie czy też przeciwnie do ruchu wskazówek zegara. Cóż doprowadziło do tego nagłego skoku rozwojowego, którego owoce przypadło nam konsumować? Okazuje się, że przyczyną całego zamieszania była zaproponowana w czasach Galileusza i Kartezjusza niezwykle prosta recepta życiowa, którą wzięła sobie do serca skromna garstka przedstawicieli naszego gatunku. Bertrand Russell receptę tę streszcza następująco:

Chcę przedłożyć czytelnikowi do życzliwego rozważenia doktrynę, która, jak się obawiam, może się wydać niesłychanie paradoksalna i wywrotowa. Według tej doktryny jest rzeczą niepożądaną wierzyć jakiemuś twierdzeniu, gdy nie ma żadnej podstawy do przypuszczenia, że jest ono prawdziwe. Muszę naturalnie przyznać, że gdyby takie mniemanie stało się powszechne, przeistoczyłoby zupełnie nasze życie społeczne i nasz ustrój polityczny; ponieważ obydwa są idealne, musi to być policzone na jego niekorzyść. Zdaję sobie również sprawę z czegoś o wiele ważniejszego, a mianowicie z tego, że wpływałoby ono na zmniejszanie się dochodów wróżbitów, bookmakerów, biskupów i innych ludzi, żyjących z irracjonalnych nadziei tych, którzy nic nie uczynili, aby uzyskać szczęście na tym lub tamtym świecie.

Przedstawiona metoda poznawania rzeczywistości sprowadza się do wątpienia i dociekliwego sprawdzania wszystkiego, co da się sprawdzić. To tyle. Żadnej wiary w opinie lub prawdy objawione. No bo w zasadzie nie da się „udowodnić" żadnego prawa przyrody. A żeby którekolwiek podważyć, wystarczy podać zaledwie jeden kontrprzykład.

Jak na przykład udowodnić, że jutro rano po raz kolejny wzejdzie Słońce? Nie wynika to przecież wcale z faktu, że w trakcie poprzednich paru tysięcy lat z rzędu Słońce co dzień rano wstawało. A wystarczyłoby zaobserwować tylko jedną taką sytuację, w której Słońce nie wzeszło o poranku, żeby prawo o codziennym wstawaniu Słońca zostało bezpowrotnie sfalsyfikowane. No i należałoby wówczas stosownie zmodyfikować prawa przyrody.

Tak właśnie działa metoda naukowa – nie poprzez udowadnianie czegokolwiek, bo się przecież nie da, a przez nieustające próby podważania tego, co aktualnie wiemy o świecie. Tak zwane „fundamentalne prawa fizyki" to po prostu zbiór twierdzeń, które póki co nie zostały empirycznie podważone. Fizycy wcale w owe prawa nie „wierzą"; lecz przyjmują je do wiadomości, jako tymczasowy model rzeczywistości, zgodny z dotychczasowymi obserwacjami. Niczego nie możemy być pewni stuprocentowo. Niektóre znane prawa przyrody wydają się bardziej prawdopodobne od innych, część z nich obowiązuje tylko w pewnym zakresie zjawisk, poza którym ulegają one załamaniu, jeszcze inne są tylko pewnymi przybliżeniami. Jest tu spora różnorodność. Tym bardziej, pod karą smoły i pierza, powinno się zabronić używania absurdalnego określenia „udowodnić naukowo" które jest nie tylko zwykłym oksymoronem, ale bezczelnym zaprzeczeniem samej idei nauki. Nauka niczego nie „udowadnia", a jedynie bada konsekwencje hipotez, które dotąd nie zostały przez nikogo skutecznie podważone. A jeśli zostały, to się nikt nie obraża, tylko wszyscy zaczynają szukać od nowa.

W matematyce mówi się co prawda o „dowodzie matematycznym" jakiegoś twierdzenia. Jednak chodzi tu tylko o badanie konsekwencji różnych założeń czy aksjomatów. A ich się już nie „udowadnia" bo nie ma jak, tylko przyjmuje za punkt wyjścia w prowadzonych rozważaniach. Założenia te mogą być kompletnie abstrakcyjne, nie mieć żadnego związku z rzeczywistością bądź nawet jej przeczyć. Nie ma to najmniejszego znaczenia dla poprawności samego „dowodu matematycznego". Zresztą praw logiki i wnioskowania też się nie udowadnia, tylko przyjmuje w formie aksjomatów wywiedzionych z naszego doświadczenia.

Receptę na ciągłe wątpienie i nie branie niczego na wiarę wziął sobie do serca najwyżej promil naszej populacji, ale i to wystarczyło, by błyskawicznie odkryć teorię ewolucji, teorię względności i mechanikę kwantową. Błyskawicznie w porównaniu z tempem rozwoju naszej cywilizacji osiąganym w trakcie poprzednich tysiącleci, gdy obowiązującą doktryną badawczą była wiara w prawdy objawione.

To, że metoda badania świata, oparta na ciągłym wątpieniu i kwestionowaniu, doprowadziła do spektakularnego sukcesu, nie jest szczególnie zaskakujące. W drugim rozdziale okaże się, co trzeba zrobić, żeby nakłonić człowieka do wiary w zupełnie dowolną, choćby najbardziej absurdalną tezę. Wiary, której człowiek ten będzie zawzięcie bronić. I to powinno wystarczyć do przekonania się, że ludzkie wierzenia nie są zbyt miarodajnym wyznacznikiem prawdy nawet w najprostszych kwestiach. Nie wspominając o fundamentalnych zagadnieniach, o których nie mamy zielonego pojęcia.

Z powodów tych gorąco zalecam pogodzenie się ze stanem faktycznym: nie jesteśmy stworzeniami nazbyt przenikliwymi. Świnia na przykład, na której wielu z nas opiera swoją dietę, jest zwierzęciem całkiem rozsądnym. Brytyjscy naukowcy badali, jak często różne zwierzęta giną rozjechane przez samochód w trakcie przechodzenia przez jezdnię, w stosunku do liczby podjętych prób. W zestawieniu najbezpieczniej zachowujących się zwierząt to właśnie świnia zajęła pierwsze miejsce. Człowiek uplasował się na miejscu czwartym.

Jeśli zapytać kogokolwiek, jaka będzie grubość kartki papieru złożonej na pół pięćdziesiąt razy (przy założeniu, że dysponujemy odpowiednio dużym arkuszem), co usłyszymy? Udzielana odpowiedź rzadko przekracza kilkanaście centymetrów. No więc jakiej grubości byłaby kartka złożona na pół pięćdziesiąt razy?

Poprawna odpowiedź to ponad sto milionów kilometrów grubości. Czyli dwie trzecie odległości od Ziemi do Słońca. Sto milionów kilometrów! Wynik ten otrzymujemy, mnożąc grubość pojedynczej kartki papieru, czyli mniej więcej 0,1 mm przez liczbę kartek otrzymaną w wyniku pięćdziesięciokrotnego składania. A ta podwaja →

się przy każdym kolejnym złożeniu arkusza na pół:

2 × 2 × 2 × ... i tak 50 razy) .. . × 2 × 0,1mm = 2^{50} × 0,1mm

Otrzymany wynik jest ogromny, bo 2 podniesione do potęgi 50 daje liczbę większą niż milion miliardów. Widać na podstawie tego przykładu, że ludzie nie do końca opanowali jeszcze sztukę mnożenia liczb przez dwa. Być może mają głos w kwestiach, w których czują się bardziej kompetentni. Na przykład w sprawach związanych ze źródłem wszechrzeczy. Uzasadniając przy tym stawiane przez siebie tezy o porządku świata swoją wiarą.

Angielska prasa nazywa grę w totolotka podatkiem od głupoty. Już nawet nie chodzi o to, jak gracze reagują, gdy zasugerować im, żeby obstawili liczby l, 2, 3, 4, 5, 6, które, nie wiedzieć czemu, „przecież" nigdy nie wypadną. Tłumaczenie, że taka kombinacja jest równie nieprawdopodobna, jak jakakolwiek inna, oczywiście do nikogo nie trafia. Wiele „systemów" gry w totolotka opartych jest na rozpowszechnionym przypuszczeniu, że prawdopodobieństwo ponownego wylosowania zestawu liczb, które raz już wcześniej padły, jest znacząco obniżone. Spotkałem się nawet z opinią, że miejsce na Ziemi, w którym dopiero co wybuchła bomba, staje się przez to bezpieczniejsze, bo prawdopodobieństwo wybuchu dwóch bomb w tym samym miejscu jest przecież znikome. Osobom dzielącym się ze mną tym spostrzeżeniem najczęściej zalecam, żeby siadały tylko na tych ławkach w parku, które zostały suto obsrane przez ptaki. Bo tam, gdzie znajduje się sto ptasich śladów, prawdopodobieństwo trafienia kolejną dwójką powinno być już mikroskopijne.

Biorąc pod rozwagę niezliczone podobne przykłady popularnych wierzeń, strategia naukowa, polegająca na ciągłym podawaniu w wątpliwość i sprawdzaniu ludzkich opinii, wydaje się, mówiąc delikatnie, wysoce uzasadniona. Na marginesie można zresztą dodać, że postawa pełna zwątpienia jest wyrazem pokory, jaką winniśmy okazywać przyrodzie w obliczu naszej spektakularnej ignorancji. W przeciwieństwie do wiary opartej na przekonaniu o czyjejś nieomylności.

My tu sobie gadu-gadu, a tymczasem ojciec i matka mechaniki kwantowej, Niels Bohr, zapytany kiedyś, czy wierzy, że podkowa, którą umieścił nad wejściem do domu, przynosi mu szczęście – odpowiedział:

Naturalnie, że nie. Ale powiedziano mi, że podkowa działa nawet wtedy, gdy się w nią nie wierzy.

W czasach studenckich postanowiłem nauczyć się czytać szybko. Ludzie czytają zazwyczaj z prędkością jakichś trzystu słów na minutę, czyli stronę tekstu w minutę. Przeciętny człowiek mówi do siebie w myślach, co czyta, dlatego nie jest w stanie czytać szybciej, niż mówi. A nikt na Ziemi nie umie mówić szybciej niż trzysta słów na minutę. Gdyby jednak pozbyć się z głowy niepotrzebnego głosu, ograniczenie to przestałoby przecież obowiązywać. I stąd rekord świata w szybkości czytania wynoszący dwadzieścia pięć tysięcy słów, czyli prawie sto stron książki na minutę.

Tak więc uczyłem się szybkiego czytania. Zależało mi na jak najszybszym dotarciu do zawartości, a samo czytanie uważałem jedynie za niedoskonałą metodę transferu tej zawartości do wewnątrz głowy. Zatem pierwszą rzeczą, której musiałem się pozbyć, był wewnętrzny głos, którym mówiłem do swoich myśli w trakcie żarcia tekstu. Czytanie jest aktywnością przestarzałą i kosztowną.

Przeciętna książka ma kilkaset tysięcy liter. Procesor w smartfonie ma kilka miliardów tranzystorów, a jest tylko sto razy droższy. Oznacza to, że każda litera w książce jest sto razy droższa niż najnowocześniejszy tranzystor. Poprzednie osiem stron kosztowało więcej niż milion tranzystorów!

Tak czy owak, wykonywałem różne ćwiczenia, które pomagają w szybszym czytaniu. Część z tych ćwiczeń służy do poszerzenia pola widzenia. Dzięki nim widzi się nie tylko to, co leży bezpośrednio przed oczami, ale także całą okolicę. Szerokie pole widzenia pomaga ograniczyć ruchy oczami, które również niepotrzebnie spowalniają czytanie.

Moje ćwiczenia wyglądały tak: na środku kartki znajdował się rząd kropek od góry do dołu, a po lewej i prawej stronie każdej kropki napisana była jakaś litera. Należało skupić wzrok na kropce i nie odrywając go, odczytywać napisane z boku litery. Nie wolno było przy tym ani na moment ruszyć oczami.

Z początku litery były dość blisko kropek i zadanie wydawało się łatwe, ale stopniowo umieszczane były coraz dalej i dalej na boki. Aż wreszcie zamieniały się w odległe układy złożone z kilku liter. Na przykład takich:

F U
A D • Z E
H I

Wykonanie każdego kolejnego ćwiczenia wymagało coraz większego wysiłku, bo litery po bokach wydawały się coraz bardziej zamazane. Pierwszą turę ćwiczeń, na której początkowo chciałem poprzestać, wykonałem dość szybko. Dlatego od razu przystąpiłem do drugiej. Po skończeniu kolejnej porcji uznałem, że mi smakuje i nie będę jeszcze przerywał. Po dwóch godzinach przerobiłem wszystkie ćwiczenia, które miałem w mojej książce z ćwiczeniami. Trochę kręciło mi się w głowie, ale wciąż nie chciało mi się kończyć.

Na środku książki, pomiędzy stronami, dopisałem więc własne kropki i patrząc na nie, zacząłem czytać litery na sąsiednich stronach. W ten sposób spędziłem kolejnych parę godzin, aż wreszcie dotarło do mnie, że jak każdej tresowanej małpie, mnie również należy się wreszcie jakiś posiłek. Mieszkałem wtedy w szarym, obdrapanym pokoju na warszawskiej Pradze.

Podniosłem wzrok znad książki i rozejrzałem się po pokoju. To co stało się ze mną tuż potem, pamiętam do dziś. Poczułem się jak żaba, która ma oczy po bokach. Wyraźnie widziałem wszystko wokół siebie, i to bez ruszania głową! Poczułem, jakbym uległ gwałtownej degradacji w hierarchii łańcucha pokarmowego, bo zwierzyna łowna, w przeciwieństwie do drapieżników, ma szeroki rozstaw oczu. Oczy umieszczone wąsko lepiej wspomagają widzenie stereoskopowe, dając lepszą percepcję głębi, ale oczy rozstawione szeroko pomagają najpełniej obserwować otoczenie. Żeby jak najszybciej dostrzec ryzyko stania się cudzym posiłkiem. I w ten sposób zapomniałem, że jestem głodny.

Stawanie się żabą było mi już znane. Gdy w podstawówce czytałem książkę Hoimara von Ditfurtha „Na początku był wodór" którą dostałem od taty, z każdym kolejnym rozdziałem czułem, że się przepoczwarzam. Dotarła do mnie trywialność mechanizmów ewolucyjnych działających poprzez dobór naturalny, natrafiłem na niezwykłe wyjaśnienie, dlaczego bez obecności Księżyca życie na Ziemi w znanej nam formie nie mogłoby powstać i nie byłoby zorzy polarnej. Oraz dlaczego kury w czasie chodzenia ruszają głowami. Było to wszystko tak nowe i oszałamiające, że podnosząc oczy znad książki i gapiąc się na Konin, znany z tego, że

„…nieodłącznym towarzyszem w nauce fizyki jest ciągła frustracja. Ucząc się od zera, zacinałem się co chwila na tym czy innym problemie, który należało samodzielnie rozwiązać. W zasadzie po dwudziestu pięciu latach niewiele się u mnie zmieniło i nadal czuję się jak małpa, próbująca dosięgnąć ręką zbyt wysoko wiszącego banana. Łaska olśnienia, żeby spróbować patykiem leżącym obok, przychodzi nie spiesznie, dopiero jak już dam się fizyce solidnie przeczołgać.…”

spośród wszystkich ówczesnych miast wojewódzkich znajdował się na drugim miejscu pod względem wykształcenia mieszkańców (od końca), czułem się trochę, jak po nadmiernej porcji ćwiczeń na poszerzenie pola widzenia.

Jakiś czas później natrafiłem na nieznane mi wyprowadzenie równań teorii względności, z którego wynikało, że nawet zwykły, niepozorny ruch jest „obrotem czasoprzestrzeni”. Byłem tak zafascynowany, że musiałem sobie obiecać, że w trakcie przechodzenia przez ulicę nie będę o tym myśleć.

Wychodzenie na zewnątrz głowy zaczęło sprawiać mi coraz więcej trudu. Czytałem jeszcze więcej, a ponieważ robiłem to w autobusach, pociągach, na przystankach, wymyślałem najróżniejsze sposoby pozwalające unikać rozkojarzenia. Któregoś razu zauważyłem na przykład, że czytanie książki do góry nogami skutecznie odcina wszystkie zewnętrzne bodźce, zmuszając mózg do nieoczekiwanego wysiłku. Zacząłem więc czytać książki do góry nogami i robiłem to do momentu, aż mój mózg zupełnie się do tego przyzwyczaił. Z tego powodu często wychodziłem na kretyna, który nie potrafi nawet właściwie trzymać książki. Szczególnie mi to nie przeszkadzało.

W końcu zamówiłem do konińskiej księgarni podręcznik fizyki, o którym słyszałem, że jest niezły, a z którego uczyli się studenci. Nie bardzo wiedziałem, jak się z czegoś takiego korzysta, nie było nawet z kim o tym pogadać, więc potraktowałem książkę jak każdą i przebrnąłem od deski do deski przez dwa opasłe tomiska Hallidaya & Resnicka. Okazało się potem, że moje podejście do nauki było dość ekscentryczne, bo podręczników fizyki nie czyta się w ten sposób. Studenci innych wydziałów dostają często do wykucia materiał kilkuset stron w ciągu trzech dni, podczas gdy studenci fizyki uczą się nieraz zawartości kilku stron podczas miesiąca.

W ten sposób dowiedziałem się jednak kolejnej ważnej rzeczy: że nieodłącznym towarzyszem w nauce fizyki jest ciągła f r u s t r a c j a . Ucząc się od zera, zacinałem się co chwila na tym czy innym problemie, który należało samodzielnie rozwiązać. W zasadzie po dwudziestu pięciu latach niewiele się u mnie zmieniło i nadal czuję się jak małpa, próbująca dosięgnąć ręką zbyt wysoko wiszącego banana. Łaska olśnienia, żeby spróbować patykiem leżącym obok, przychodzi nie spiesznie, dopiero jak już dam się fizyce solidnie przeczołgać.

A jednak jakimś cudem stałem się psychopatą i wszystko inne przestało mnie obchodzić. Parę rozdziałów dalej wyjaśnię dlaczego. Dokupiłem jeszcze zbiór 1500 zadań Kruczka i rozwiązałem wszystkie w czasie dwutygodniowych ferii zimowych, nie robiąc w zasadzie niczego poza tym. Starałem się trzymać dzienną średnią na poziomie 100 zadań. A skoro już miałem okazję, postanowiłem przy tym przeprowadzić na sobie różne eksperymenty. Ciekawiło mnie na przykład, jak długo będę w stanie wytrzymać rozwiązywanie zadań bez snu ani jakiegokolwiek odpoczynku. Po 48 godzinach, kiedy pojawiły się pierwsze halucynacje, uznałem, że rozsądnie będzie przerwać eksperyment.

Fizyka odkleiła mnie od reszty świata. Aby nieco naświetlić niektóre przyczyny, napiszę o przykładzie, który mocno podziałał na moją wyobraźnię. W pierwszej pracy Einsteina o teorii względności była mowa o tym, że czas na zegarze, który porusza się z dużą prędkością, płynie wolniej, a rakieta poruszająca się z dużą prędkością skraca się wzdłuż kierunku, w którym leci. Efekty te robią się bardziej zauważalne, gdy ruch zbliża się do prędkości światła. Światła, które w ciągu sekundy mogłoby okrążyć Ziemię siedmiokrotnie. Doświadczane na co dzień prędkości są dużo mniejsze, więc nie dostrzegamy tych dziwacznych efektów, ale w skalach kosmicznych sprawa robi się poważna i zegary n a p r a w d ę chodzą wolniej! Co dziwniejsze, ani poruszający się zegar, ani lecąca szybko rakieta, nie mają zielonego pojęcia, że podlegają spowolnieniu upływu czasu albo skróceniu długości. No bo przecież z ich punktu widzenia one same spoczywają, a cały świat porusza się w przeciwną stronę. I według nich to wszystkie inne zegary, a nie one same, chodzą wolniej, a cały poruszający się świat ulega skróceniu. Oto pozornie paradoksalna konsekwencja względności ruchu.

Innym ciekawym efektem, o którym pisał Einstein, jest względność równoczesności. Wyobraźmy sobie dwa zdarzenia zachodzące w tej samej chwili, ale w różnych miejscach. Niech na przykład pierwszym zdarzeniem będzie złożenie jajka przez kurę we wsi pod Koninem, a drugim zdarzeniem wyklucie się kurczaka z innego jajka we wsi pod Warszawą. Przyjmijmy, że oba te zdarzenia zaszły równocześnie: gdzieś złożone zostało jajko, a gdzie indziej z innego jajka wykluła się kura. Okazuje się, że jeśli będziemy obserwować świat z poruszającego się pojazdu, to jedno z tych zdarzeń zajdzie jako pierwsze, a drugie dopiero po chwili. Kolejność zależy od kierunku, w którym będziemy jechać. Jeśli ktoś jadący w prawo stwierdzi, że najpierw było konińskie jajko, a potem wykluła się warszawska kura, to ktoś jadący w lewo uzna, że najpierw była kura, a dopiero później jajko. No i według teorii względności żaden z punktów widzenia nie jest lepszy od pozostałych. Żeby było jasne: nie chodzi o to, że informacje o zdarzeniach docierają do nas z opóźnieniem, które zależy od tego, gdzie się znajdujemy. Nic z tych rzeczy. Chodzi o to, że pojęcie „t e r a z" zmienia swoje znaczenie, gdy zaczynamy się ruszać. „Teraz" jest względne.

Jeśli komuś nie chce się wierzyć, to zapewne dlatego, że w większości codziennych sytuacji te efekty są tak niewielkie, że nie da się ich zauważyć gołym okiem. Jesteśmy przyzwyczajeni do niewielkich skal. Użyteczne jednostki odległości to „stopa", „łokieć" albo „ryczenie wołu" (czyli odległość, z której jeszcze słychać ryczącą rogaciznę). Natomiast pojęcie teraźniejszości zaczyna ulegać zmianie dopiero w trakcie ruchu z prędkością zbliżoną do prędkości światła. Albo w skalach kosmicznych, bo wówczas prędkości nie muszą być duże. Dla przykładu, jeśli z prędkością starej baby na rowerze pojedziemy w stronę gwiazdy oddalonej od nas o sto lat świetlnych, to spowodujemy, że stanie się ona starsza o mniej więcej minutę (gwiazda, nie baba). A gdy pojedziemy w przeciwnym kierunku, gwiazda o minutę odmłodnieje. Takie są wnioski z teorii względności.

Nie ma się co dziwić, że wielu fizyków zareagowało na pracę Einsteina wrogo. Powstała nawet książka „Stu autorów przeciwko Einsteinowi", w której na sto sposobów „udowadniano" dlaczego twórca teorii względności gada bzdury. Gdy ten dowiedział się o owej publikacji, słusznie zauważył, że gdyby wszyscy owi naukowcy mieli rację, wystarczyłby jeden.

Minęło trochę czasu, zanim eksperymentalnie potwierdzono przewidywania teorii względności, a starsze pokolenie fizyków nieco ochłonęło lub po prostu wymarło. Einstein zaś stał się celebrytą. Ale nie o tym chciałem pisać.

> „...Wszystkie przewidywania teorii względności zostały potwierdzone w niezliczonych eksperymentach, a mimo to nasza reakcja na nią, nawet po ponad stu latach od odkrycia, może rodzić najróżniejsze powikłania umysłowe..."

Można na wiele różnych sposobów zrozumieć źródło wszystkich tych nieintuicyjnych konsekwencji teorii względności. W większości są to sposoby dość zawiłe, a jeden z nich podał zresztą sam Einstein w swojej pracy. Lecz czasem można odnaleźć drogę prowadzącą do wyniku, która okazuje się jeszcze ciekawsza niż sam wynik.

Coś takiego zdarzyło się w 1907 roku, czyli tuż po opublikowaniu pierwszej pracy Einsteina, i właśnie t o odkrycie tak podziałało mi na wyobraźnię. Tego roku dawny nauczyciel Einsteina, Herman Minkowski, przedstawił swój sprytny nowy sposób patrzenia na teorię względności. Cóż takiego zauważył Minkowski? Równania otrzymane przez Einsteina, z których wynikają wszystkie zaskakujące przewidywania teorii względności, opisują, jak się zmienia czas i położenie różnych zdarzeń, gdy obserwujemy je, będąc w ruchu. Gdyby ktoś był ciekawy, równania wyglądają tak:

$$x' = \frac{x - Vt}{\sqrt{1 - V^2}} \qquad t' = \frac{t - Vx}{\sqrt{1 - V^2}}$$

Stephen Hawking stwierdził co prawda, że umieszczanie jakiegokolwiek matematycznego równania w popularnej książce zmniejsza liczbę czytelników o połowę, ale Hawking stwierdził przy innej okazji, że fizyka zostanie ukończona do końca XX wieku, więc może nie ma się czym przejmować.

Powróćmy do Minkowskiego. Zwrócił on uwagę, że powyższe równania do złudzenia przypominają inne równania, znane z elementarnej szkolnej geometrii. A mianowicie równania opisujące obrót układu współrzędnych o pewien kąt, które wyglądają tak:

$$x' = \frac{x - At}{\sqrt{1 + A^2}} \qquad t' = \frac{t + Ax}{\sqrt{1 + A^2}}$$

Wystarczy zerknąć i porównać, podobieństwo jest uderzające! Różnica między równaniami jest tylko taka, że prędkość V obserwatora (zapisana jako ułamek prędkości światła) w równaniach teorii względności jest zastąpiona tangensem A kąta, o który obrócony jest układ współrzędnych. Jest też we wzorach inna subtelna modyfikacja: zamiast znaku minus w trzech miejscach znajduje się znak plus. Ta obserwacja nie dawała Minkowskiemu spokoju. Czyżby z matematycznego punktu widzenia ruch był po prostu pewnym rodzajem ... obrotu? Ale niby obrotem c z e g o ?! Pytanie to wymagało z pewnością oderwania się intelektualnie od żucia gumy.

Problem postawiony został dość abstrakcyjnie, więc jedynymi obiektami, które pojawiły się w rozważaniach i mogły w trakcie ruchu podlegać obrotowi, były czas t i przestrzeń x występujące w równaniach! Minkowski wprowadził więc pojęcie *czasoprzestrzeni* i ogłosił, że to, co odczuwamy jako ruch, może być w istocie ujęte matematycznie, jako pewien specyficzny obrót owej czasoprzestrzeni. Ze względu na różnicę w znakach pojawiających się w równaniach, ten dość nietypowy obrót nazywany jest *obrotem hiperbolicznym*. Ale ma on wiele właściwości najzwyklejszych obrotów znanych z codziennego życia. A co jest najciekawsze w tej obserwacji poza tym, że brzmi ona dość nieprawdopodobnie? Otóż to, że jeśli zacisnąć zęby i przyjąć ją za punkt wyjścia, to prawie wszystkie wnioski płynące z teorii względności stają się wręcz banalne!

Zapomnijmy na moment o teorii względności i wyobraźmy sobie, że trzymamy w ręku kij. Każdy kij ma dwa końce (z wyjątkiem procy). Czy oba końce kija są równo oddalone? Chwila na zastanowienie. No cóż, nasze pytanie nie ma sensu. To tak, jak by zapytać: czym się różni wróbelek? Aby nadać pytaniu sens, trzeba najpierw doprecyzować: od czego mają być równo oddalone końce naszego kija? No bo mogą one być równo oddalone od czyjegoś lewego oka, ale wówczas od prawej pięty już raczej nie. Wystarczy obrócić kij i sytuacja ulega zmianie.

Jeżeli ruch rzeczywiście jest pewnym obrotem czasoprzestrzeni, to pytanie, czy dwa zdarzenia są równoczesne, czy nie, jest tak samo pozbawione sensu, jak pytanie, czy dwa końce kija są równo oddalone! Musimy najpierw sprecyzować, w z g l ę d e m k t ó r e g o obserwatora.

Według jednego obserwatora najpierw było jajko, a potem kura, a dla innego, poruszającego się obserwatora, najpierw była kura, a dopiero potem jajko. A dla jeszcze innego kura i jajko pojawiły się na świecie równocześnie. I nie ma tu żadnej sprzeczności. Podobnie jak nie ma sprzeczności pomiędzy tym, że według jednego obserwatora pierwszy koniec kija jest bliżej, a z punktu widzenia innego obserwatora, według którego kij jest obrócony, to drugi koniec może być bliżej. Nie ma nawet sensu pytać, który z obserwatorów ma rację. Istota teorii względności jest naprawdę aż tak prosta.

Trzymajmy się myślowo kija. Jego długość też zależy od punktu widzenia. Fotograf robiący kijowi zdjęcie pod pewnym kątem stwierdzi, że na fotografii kij wyszedł krótszy niż wówczas, gdy kij jest równoległy do matrycy aparatu. I takim samym zmianom ulega długość czasu zmierzonego na zegarze, względem którego się ruszamy! Zgodnie ze sformułowaniem Minkowskiego, ruszając się, obracamy całą czasoprzestrzeń, a więc odcinki czasu odmierzane przez poruszający się względem nas zegar ulegną skróceniu. Co oznacza, że zegar po prostu będzie chodzić wolniej. Jest to wniosek, który Einstein skrupulatnie wydedukował ze swoich równań, lecz my, dzięki temu, że interpretujemy ruch jako obrót czasoprzestrzeni, możemy się tego spodziewać b e z wykonywania jakichkolwiek obliczeń. Coś pięknego!

Ponieważ w wyniku ruchu obraca się cała czasoprzestrzeń, czyli nie tylko oś czasu, ale również oś przestrzeni, podobnie skrócenie musi dotyczyć też fizycznej długości ruchomych ciał. Gdy więc jedzie rower, to będzie on spłaszczony wzdłuż kierunku, w którym się porusza. Niecodzienny ów wniosek polecam do rozważenia wszystkich, którzy czytając te słowa, podróżują akurat dowolnym środkiem lokomocji wyposażonym w okna.

Pomysł Minkowskiego jest może łatwy do sformułowania, ale wymaga wzmożonego trawienia organem, który nosimy 15 cm z tyłu nosa. Nie potrafię myśleć bardzo szybko i pamiętam, że po tym, jak przeczytałem o obrotach czasoprzestrzeni, odzyskanie równowagi psychicznej zajęło mi ze dwa tygodnie. Tym bardziej że obrotu czasoprzestrzeni, o którym mądra książka pisała dość lekko, za nic nie potrafiłem sobie wyobrazić. Pocieszałem się, że moja percepcja przestrzenna ewoluowała głównie po to, żebym jak najskuteczniej rozstrzygał, które spośród okolicznych organizmów wielokomórkowych są jadalne, a które wręcz przeciwnie i najwyżej w tym mogę być dobry. I nie po to moi niedoszli przodkowie ginęli w wyniku selekcji naturalnej, żebym intuicyjnie pojmował zjawiska zachodzące, gdy poruszam się z prędkością zbliżoną do prędkości światła. A dopiero wówczas ów obrót hiperboliczny, o którym pisał Minkowski, staje się zauważalny. Bo dla niewielkich prędkości, z którymi mamy do czynienia na co dzień, ruch w dobrym przybliżeniu wydaje się po prostu niewinną zmianą położenia obiektów w czasie.

Okazało się jednak, że do wszystkiego można się przyzwyczaić. Więc i ja w końcu się przyzwyczaiłem, że ruch po prostu j e s t niewyobrażalnym obrotem czasoprzestrzeni. I zapomniałem, że jest to coś dziwnego. Było to dla mnie o tyle łatwiejsze, że teoria względności do swojego zrozumienia nie wymaga nawet minimum wiedzy matematycznej, której wówczas nie posiadałem. Pomysł Minkowskiego, jeśli już się z nim pogodzić, pozwala pojąć najistotniejsze konsekwencje teorii względności, nawet dysponując niewiedzą wręcz encyklopedyczną.

Moja prywatna czasoprzestrzeń również uległa obróceniu: zdecydowałem, że będę studiować fizykę, a gdy ktoś mnie pytał, co mam zamiar robić po takich studiach, odpowiadałem, że będę pracować fizycznie.

Wszystkie przewidywania teorii względności zostały potwierdzone w niezliczonych eksperymentach, a mimo to nasza reakcja na nią, nawet po ponad stu latach od odkrycia, może rodzić najróżniejsze powikłania umysłowe. W moim przypadku były to niemal łzy radości. U niektórych zdarzają się jednak łzy smutku, ale najczęściej są to po prostu łzy konsternacji...

***Nota wydawnicza:**

Tekst jest fragmentem książki Andrzeja Dragana pod ty samym tytułem „Kwantechizm czyli klatka na ludzi", Fabuła Fraza Sp. z o.o. , Warszawa 2019

{Andrzej Dragan (rocznik 1978) zajmuje się łączeniem ogólnej teorii względności z teorią kwantową na Uniwersytecie Warszawskim. Profesor wizytujący na Uniwersytecie w Singapurze, laureat nagrody „Polityki", Ministra Edukacji Narodowej, Fundacji na rzecz Nauki Polskiej, European Science Foundation i Polskiego Towarzystwa Fizycznego. Fotograf Roku brytyjskiego magazynu „Digital Camera", nominowany do Złotego Lwa na festiwalu reklamowym w Cannes, zdobywca Best in Show brytyjskiego magazynu „Creative Review" i Złotego Miecza KTR za debiutancki film (Dylatacja czasu), nominowany do nagrody Yacha za debiutancki teledysk (Behemoth). Od 2018 roku współpracuje z Centrum Studiów Zaawansowanych w ramach Uczelnianej Oferty Dydaktycznej PW oraz cykli Dysputy Pitagorejskie i Spotkania Otwartych Umysłów}

O innowacyjności

List otwarty Przewodniczącego Rady Fundacji na rzecz Nauki Polskiej, profesora Leona Gradonia

Postawa twórcy powinna być kształtowana na wszystkich poziomach edukacji i w programach stymulujących kreatywność. Są bowiem dwa fundamentalne cele nauki – rozumieć świat i oddziaływać na niego. Sposób oddziaływania i jego skutki są różne na różnych poziomach aktywności twórczej. Jedną z nich jest innowacyjność związana z komercjalizacją badań naukowych.

Fundacja na rzecz Nauki Polskiej w aktywny sposób włącza się w stymulowanie wysokiej jakości badań, w tym takich, które służą wspieraniu transferu osiągnięć naukowych do praktyki gospodarczej. FNP była prekursorem w tworzeniu programów wspierających badania proinnowacyjne. Już w latach 90. dofinansowywała projekty o charakterze rynkowym, oparte na nowych technologiach. Rezultaty oddziaływania tych programów w przestrzeni biznesowej były skromne, głównie ze względu na warunki zewnętrzne, ale także jakość proponowanych rozwiązań w projektach składanych do pierwszych programów tego typu. W kolejnych latach Fundacja proponowała coraz bardziej dopracowane formy wspierania innowacyjnych wdrożeń.

Obecnie, ze względu na rozwój technologiczny oraz możliwości tworzenia małych i średnich przedsiębiorstw, istnieje większe zapotrzebowanie na wdrażanie nowych technologii do gospodarki. W ostatnich latach Fundacja uruchomiła nowe programy ukierunkowane na wdrożenia (FIRST TEAM, TEAM i przede wszystkim – TEAM-TECH.) Zasadą przyznawania grantów jest tu tryb konkursowy, ocena wniosków metodą *peer-review*, wymóg doskonałości naukowej i ścisłe procedury rozliczania wyników. Rezultaty takiego sposobu postępowania są już widoczne. Programy FNP wyłaniają prawdziwych liderów naukowych. Taki sposób postępowania może być wykorzystany w programach grantowych innych instytucji na etapie reformy nauki.

Nauka z definicji i swojego charakteru jest synonimem innowacyjności. Część aktywności naukowej, na pewnym etapie dojrzałości technologicznej, ma bezpośrednie przełożenie na gospodarkę.

Innowacyjność jest rodzajem aktywności twórczej, która w sposób naturalny jest włączona w system gospodarczy nowoczesnego państwa i staje się siłą napędową jego rozwoju. Jednocześnie powoduje ona zastąpienie imitacyjnego modelu gospodarczego, powielającego w znacznej mierze produkcję krajów bardziej zaawansowanych technologicznie.

> „Umiejętności dopotąd są jeszcze próżnym wynalazkiem, może czczym tylko rozumu wywodem albo próżniactwa zabawą, dopokąd nie są stosowane do użytku narodów”
>
> *| Stanisław Staszic |*

W jaki sposób nauka powinna być wykorzystywana jako element takiego systemu?

Z doświadczenia liderów technologicznych wynika, że sama nauka, pozbawiona dodatkowych elementów wspomagających ten obszar aktywności, nie staje się źródłem innowacyjności. Należy budować ją na kilku filarach, m.in.:

- strategii państwa wyznaczającej perspektywę i drogę do osiągnięcia dobrobytu,
- edukacji na wszystkich szczeblach kształcenia,
- systemie prężnych przedsiębiorstw nastawionych na wdrażanie i rozwijanie nowych technologii.

Szczególną rolę trzeba przypisać tu edukacji, w toku której należy rozwijać postawy kreatywnego myślenia i działania. Pożywką dla kreatywności, we wszelkich jej aspektach, jest wolność wyzwalająca niezależność myślenia, pozbawiona ograniczeń wynikających z rutyny, schematów i przyzwyczajeń.

W systemie edukacji w szczególny sposób należy uwzględnić ważną składową, którą jest wykształcenie akademickie (łączące nauczanie oraz badania naukowe prowadzone w warunkach konkurencji i rywalizacji). Tak zorganizowany system szkolnictwa wyższego pozwala tworzyć elity, których przedstawiciele zasilą m.in. firmy, przyczyniając się do wzrostu możliwości rywalizacji naszego kraju na światowych rynkach w warunkach dużej konkurencyjności. Przy wspomnianym sposobie działania, kiedy przekroczony zostanie pewien poziom wartości PKB, zaczyna się istotny wzrost nakładów na badania, również te o charakterze podstawowym. Dopóki w Polsce nie zostanie wytworzony podobny mechanizm, pozostaniemy w zaklętym kręgu, w którym głównym żądaniem uzasadniającym podniesienie jakości badań będzie wzrost wartości środków budżetowych przeznaczonych na naukę.

Klasycznym przykładem gospodarki innowacyjnej są Stany Zjednoczone. Tam powstaje najwięcej nowych technologii, a uniwersytety są w czołówce światowej wg dowolnych kryteriów rankingowych. Podobny model działania przyjęły „azjatyckie tygrysy”. W państwach tych wspaniale rozwijają się gospodarki, a uniwersytety tych państw zajmują wysokie pozycje w rankingach jakości badań i kształcenia studentów.

Kraje europejskie nie osiągnęły podobnego poziomu aktywności innowacyjnej. Najlepiej w Europie wypadają kraje skandynawskie, a Polska zajmuje 24 miejsce w tej klasyfikacji. Obserwując aktywność naukową i tematykę badań podejmowanych w naszym kraju, widać proste naśladownictwo i często zaściankowość. System wyrzuca na zewnątrz tych bardziej ambitnych i przedsiębiorczych, którzy muszą szukać swojego miejsca poza nauką lub w zagranicznych laboratoriach badawczych.

W teorii Adama Smitha, na którą od ponad dwustu lat powołuje się wielu ekonomistów, społeczeństwo stanowi zbiór racjonalnych jednostek, które w swoim działaniu kierują się zasadą własnej korzyści (*homo economicus*). Owa zasada racjonalności w działalności gospodarczej, ujawniająca się w stanie „zupełnej wolności", przenoszona jest zwykle również na działalność zbiorowości.

Postawa *homo economicus* jest również ważnym elementem wciągającym jednostki w działalność innowacyjną, w której łączy się wiedzę, doświadczenie techniczne i intuicję z przedsiębiorczością. Taki rodzaj aktywności należy wyzwolić. Że tak się w sposób naturalny staje, pokazuje nam tzw. „Ustawa Wilczka" (ustawa z 23 grudnia 1988 r. o działalności gospodarczej), wprowadzona w schyłkowym okresie PRL.

Duża swoboda i wolność ekonomiczna, będąca podstawowym elementem tej ustawy, zaowocowały wybuchem intensywnej aktywności twórczej i rozwoju przedsiębiorczości prywatnej. Powstało wiele firm wdrażających nowe rozwiązania techniczne na poziomie technologicznym, który umożliwiała sytuacja ekonomiczna kraju w tamtym czasie. Paradoksalnie swoboda ta została ograniczona formalnymi uregulowaniami wprowadzonymi w okresie transformacji ustrojowej. Dodatkowo, niekorzystnie przeprowadzona prywatyzacja doprowadziła do utraty bazy materialnej, na której można było rozwijać przedsiębiorczość prywatną. Do Polski weszło wiele firm zagranicznych, które wprowadziły własne technologie i reguły gry związane z patentowaniem i ochroną własności intelektualnej. Wykształceni inżynierowie dostawali w nich głównie funkcje *product manager*, co sprowadzało ich do roli sprzedawców. Fachowa wiedza, którą dysponowali oraz ich pomysłowość, nie były w pełni wykorzystywane. Wielu aktywnych i kreatywnych absolwentów szkół wyższych postanowiło szukać możliwości wykazania swojej mocy twórczej w pracy za granicą. Dziesiątki tysięcy młodych zdolnych zasiliło swoją pracą gospodarki krajów m.in. zachodniej Europy. W kraju następował spadek aktywności badawczej. Polska w rankingu poziomu naukowego obniżała swoją pozycję. Ciekawą analizę tego zjawiska przedstawił prof. Andrzej K. Wróblewski w artykule opublikowanym w tygodniku WPROST (grudzień, 1998).

Próba zmiany tej sytuacji nastąpiła po wstąpieniu Polski do Unii Europejskiej. Dotyczyło to również obszaru, gdzie wprowadzono mechanizmy wspomagania nauki przez system grantowy, z założenia wyzwalający rywalizację w pozyskiwaniu środków na badania. Na miejsce istniejącego w latach 1991-2005 KBN-u, system grantowy realizowany jest głównie przez NCN, NCBiR, a także Ministerstwo Nauki i Szkolnictwa Wyższego, pozarządową Fundację na rzecz Nauki Polskiej i inne agendy. W ramach m.in. funduszy europejskich powstały centra badawcze, uczelniane lub środowiskowe, z nowoczesną infrastrukturą techniczną i naukową. Z założenia miały one spełniać rolę platform transferu technologii. Niestety, jak dotąd nie widać aktywności tych instytucji.

Rozwój innowacyjności jest pobudzany głównie przez umiejętny transfer rozwiązań powstałych w wyniku badań naukowych, do biznesu. Źródeł innowacyjności można poszukiwać w wynikach badań podstawowych i stosowanych. Większość programów grantowych NCBiR wymusza, już na etapie składania wniosków, tworzenie konsorcjów, które w swoim składzie mają przedsiębiorstwa i jednostki badawcze.

Pewne światło na to, jak wykorzystywane są środki na komercjalizację wiedzy (rozumianą jako całokształt działań związanych z jej przekształcaniem na nowe produkty, technologie i rozwiązania organizacyjne) rzuca raport NIK (nr 227/2015/P/15/027/KNO) będący informacją o kontroli pt. „Komercjalizacja badań naukowych". Głównym celem kontroli była ocena skuteczności wsparcia komercjalizacji wyników badań w obszarze zwiększenia innowacyjności gospodarki, a motywacją – uniknięcie ryzyka wystąpienia zjawiska pułapki średniego dochodu jako konsekwencji wsparcia innowacyjności. Ministerstwo Nauki i Szkolnictwa Wyższego oraz NCBiR otrzymały w latach 2011-2015 na realizację programu wsparcia komercjalizacji badań naukowych kwotę ok. 3,4 mld złotych. Raport szczegółowo omawia sposób zagospodarowania tych pieniędzy. W konkluzji, w zbyt krótkiej perspektywie czasowej dla możliwości pełnej oceny, NIK widzi następujące bariery w osiągnięciu celu w tych przedsięwzięciach:

- → niskie zainteresowanie przedsiębiorców działaniem innowacyjnym i niewielki strumień środków z firm do jednostek naukowych,
- → rozproszona tematyka badawcza i brak określenia polskiej specjalizacji w obszarze nauki,
- → nieskuteczny sposób nadzoru nad realizacją programów.

W ocenie NIK najkorzystniej wypadły te programy, których celem był rozwój aktywności młodych badaczy (np. TOP500 Innovator i Broker Innowacyjności). W ocenie większości programów wskazywano na nieprzestrzeganie terminowości kończenia prac i wady projektowe uniemożliwiające weryfikację uzyskanych rezultatów. Znaczna liczba zgłoszeń wynalazków do Urzędu Patentowego, będąca rezultatem realizowanych projektów, wygasała po upływie pierwszego roku ochrony. Wskazuje to na niski potencjał ich komercjalizacji i pozwala przypuszczać, że zgłoszenia były jedynie elementem formalnego rozliczenia się zespołów z wykonywanych zadań.

Jedną z przyczyn braku aktywności twórczej i przedsiębiorczości jest system oceny pracy badawczej sprowadzający się do zdobywania punktów. Takiej formie oceny podlegają również zgłoszenia patentowe (sic!). Główną motywacją uprawiania nauki jest chęć poznania prawdy, odkrycia czegoś nowego czy też skonstruowania użytecznego urządzenia. Obowiązująca obecnie ocena parametryczna osłabia aktywność twórczą. Czy 200 punktów za jeden świetny artykuł będący podsumowaniem wartościowej pracy jest równoważne 200 punktom uzyskanym z publikacji 5 przyczynkarskich artykułów „wycenianych" na 40 punktów? Ile warte jest zgłoszenie patentowe, puste pod względem potencjalnej możliwości aplikacji, skoro znika z procedury ochrony praw po roku od daty zgłoszenia? Autorzy projektów starają się rozliczyć formalnie swoje granty przez takie właśnie zgłoszenia. Patent ma wartość wtedy, gdy wyzwala zainteresowanie rynkowe. Ocenie powinny podlegać nie zgłoszenia patentowe, ale przyznane patenty poparte potencjalnym zainteresowaniem wdrożeniowym, zakupem licencji czy też zakupem praw do stosowania rozwiązań przez przedsiębiorców.

Dla wyzwolenia ducha kreatywności i innowacyjności potrzebne jest stworzenie mechanizmów wyłaniania liderów, w których tkwi cecha wojownika. Takie cechy ma prawdziwy inżynier, który podejmuje wyzwanie przejścia od prac badawczych do prac wdrożeniowych i osobiste ryzyko wzięcia

→

„Dla wyzwolenia ducha kreatywności i innowacyjności potrzebne jest stworzenie mechanizmów wyłaniania liderów, w których tkwi cecha wojownika"

odpowiedzialności za prawidłowość proponowanych rozwiązań. Dotyczy to również naukowców, których wyniki badań podstawowych przeniesione do biznesu stają się jedną z najskuteczniejszych form rozwoju gospodarki.

Jak wyzwolić taki rodzaj aktywności? Jednym z elementów jest program nauczania na wszystkich szczeblach edukacji oraz kryteria oceny wniosków i weryfikacja wyników badań naukowych realizowanych w ramach projektów badawczych finansowanych przez agendy grantowe. System organizacji nauki i nauczania powinien być tak budowany, aby wprowadzane do niego pieniądze nie były marnotrawione. Na obecnym etapie wdrażania Ustawy 2.0 należy zwrócić na te elementy szczególną uwagę i odpowiedzialnie monitorować realizację tych celów. W wyniku wprowadzonych zmian legislacyjnych powinny powstać mechanizmy pobudzające do aktywności projakościowych. W konsekwencji ich działania nastąpi naturalna stratyfikacja jakości szkół wyższych, wzmacniająca elementy rywalizacji, które powinny być siłą napędową ich ciągłego rozwoju. Istotną rolę mogą odegrać Rady Szkół Wyższych, w skład których powinny wchodzić osoby z dużym doświadczeniem biznesowym. Ich zaangażowanie powinno wyzwolić aktywność proinnowacyjną w uczelniach.

Reforma nauki w obszarze pobudzania innowacyjności musi być skoordynowana z rządowym programem działań proinnowacyjnych. Ważnym czynnikiem w takim programie jest stworzenie warunków sprzyjających przedsiębiorczości i zdefiniowanie strategicznych programów, które mogą stać się polską specjalnością. Synchronizacja takich działań może spowodować uruchomienie skutecznych mechanizmów włączenia innowacyjności w rozwój ekonomiczny kraju. Dla wyzwolenia postaw kreatywnych potrzebna jest pełna wolność gospodarcza, bez której nie powstaną wartości niezbędne dla innowacji.

Musimy pamiętać, że innowacyjność to nie jest gotowy produkt, który chcemy wytworzyć, ale jest to proces prowadzący do produktu. Szczególną rolę odgrywa w nim edukacja na wszystkich poziomach kształcenia, jak to już zaznaczyłem wyżej. W programach nauczania i sposobie ich realizacji powinien znaleźć się pierwiastek pobudzania kreatywności jako jedno z ważnych kryteriów oceny wyników pracy ucznia. Z takiego sposobu działania wyłaniają się liderzy – przyszli twórcy. Szczególnie dotyczy to kształcenia na poziomie akademickim, gdzie w ten proces włączane są badania naukowe. Należy zdawać sobie sprawę, że innowacyjność jest synonimem nauki, czyli dokonań nowatorskich. Jeśli nauka nie jest nowatorska, staje się jedynie (sic!) wiedzą.

Jednym ze sposobów powstania i uruchomienia kapitału ludzkiego do działań proinnowacyjnych jest zdefiniowanie spójnego programu strategicznego skrojonego na bieżące możliwości kraju. Na bazie takiego programu powinna rozwijać się aktywność gospodarcza, prowadząca do rozwoju i wzmocnienia pozycji klasy średniej, której przedsiębiorczość jest głównym źródłem wzrostu PKB w każdej liberalnej gospodarce. Dzięki republikańskim postawom możemy uwierzyć w nasze możliwości twórcze i zerwać z postawami charakterystycznymi dla krajów postkolonialnych.

Opracowanie spójnego programu strategicznego, łączącego aktywność naukową z dobrze zdefiniowanym programem rozwoju gospodarczego, stanie się kołem napędowym wyzwalającym inicjatywę innowacyjności. Jest to potrzeba chwili.

Profesor Leon Gradoń

***Nota wydawnicza:**
Przedruk z Raportu Rocznego 2018 Fundacji na rzecz Nauki Polskiej, za zgodą wydawcy.

{**Profesor Leon Gradoń** - specjalista w zakresie inżynierii biomedycznej oraz inżynierii chemicznej i procesowej. Absolwent inżynierii chemicznej na Politechnice Warszawskiej (1969) oraz matematyki na Uniwersytecie Warszawskim (1975). W 1990 otrzymał tytuł profesora nauk technicznych. Zawodowo związany z Politechniką Warszawską. Od 1999 do 2005 pełnił funkcję dziekana Wydziału Inżynierii Chemicznej i Procesowej. Przewodniczył Radzie Naukowej Centralnego Instytutu Ochrony Pracy. Był stypendystą Fundacji Fulbrighta w Cincinnati, wykładał jako *visiting professor* na uczelniach w USA, Japonii, Austrii, Holandii i Szwecji. Jest autorem lub współautorem licznych publikacji naukowych, w tym kilkunastu monografii i podręczników akademickich, a także kilkudziesięciu patentów oraz wdrożeń przemysłowych. W 2006 otrzymał Nagrodę Fundacji na rzecz Nauki Polskiej za „wyjaśnienie podstawowych procesów transportu w układach dwufazowych i ich wykorzystanie do opracowania nowych konstrukcji filtrów wgłębnych". Był wiceprzewodniczącym Komitetu Inżynierii Chemicznej i Procesowej Polskiej Akademii Nauk, powoływany w skład rad i komitetów doradczych różnych organizacji i instytucji (m.in. Narodowego Centrum Badań i Rozwoju) oraz czasopism naukowych. W 2016 roku został wybrany na przewodniczącego Rady Fundacji na rzecz Nauki Polskiej}

Sieci neuronowe - wprowadzenie

Władysław Homenda, Wydział Matematyki i Nauk Informacyjnych PW

Uwagi wstępne

Niniejsza praca została przygotowana na podstawie książki [3]. W zamierzeniu ma przedstawić idee sieci neuronowych. Z tej przesłanki wynika zakres tekstów źródłowych. Są to teksty z okresu powstania idei sieci neuronowych i formowania pierwszych realizacji takich sieci. Są też odwołania do późniejszych pomysłów istotnie uzupełniających pierwotną idee oraz do podsumowań.

Sieci neuronowe są bodaj najbardziej rozwiniętym narzędziem stosowanym w szeroko pojętej informatyce. Początków sieci neuronowych można się doszukać w pierwszej połowie lat czterdziestych ubiegłego wieku w pracy McCullocha i Pittsa [8], w której zaproponowali model neuronu w postaci binarnego automatu ze stochastycznym algorytmem przełączeń stanów. Pod koniec lat czterdziestych w pracy Hebba [2] została opublikowana koncepcja uczenia skojarzeniowego. Druga połowa lat pięćdziesiątych i początek lat sześćdziesiątych to okres prac nad perceptronem, którego idee podał Rosenblatt [10, 11]. Został podany dowód, że uczenie perceptronu na liniowo rozdzielnych zbiorach wzorców jest zbieżne. Ta obiecującą idea sztucznego neuronu okazała się mało przydatna w zastosowaniach praktycznych. W rezultacie pod koniec lat sześćdziesiątych idea sztucznych sieci neuronowych znalazła się w ogniu krytyki, począwszy od pracy Minsky'ego i Paperta [9], a skończywszy na pracy Simona [13], w której poświecił perceptronowi fragment zatytułowany *Birth and Death of a Myth*. Sytuacja zmieniła się w drugiej połowie lat osiemdziesiątych, kiedy Rumelhart, Hinton i Williams [12] przedstawili idee perceptronu wielowarstwowego i metodę uczenia przez wsteczną propagacje błędów. Mimo że nie ma formalnego dowodu zbieżności procesu uczenia perceptronu wielowarstwowego, to w praktyce uzyskuje się niezłe rezultaty, które umożliwiły realne zastosowania perceptronu wielowarstwowego. Uwaga: wspomniany tutaj perceptron nie będzie szczegółowo omówiony w niniejszej pracy. Można przyjąć, że perceptron jest siecią neuronową z progową bipolarną funkcją aktywacji, szczegóły są w [3].

Sieć jednowarstwowa – liniowe rozdzielenie

Przez pojęcie sieci jednowarstwowej rozumiemy model sztucznej sieci neuronowej z jedną warstwą połączeń. Będzie to oczywiście warstwa połączeń sygnałów wejściowych z neuronami warstwy wyjściowej. Przykład sieci jednowarstwowej z n sygnałami wejściowymi i jednym sygnałem wyjściowym jest przedstawiony na rysunku 2. Natomiast na rysunku 3 przedstawiono sieć jednowarstwową o c wyjściach.

Przypadek dwuklasowy

Załóżmy, że naszym zadaniem jest klasyfikacja elementów opisanych wektorami cech z przestrzeni $\mathbb{R}^n$ do jednej z dwóch klas C_1 i C_2. Załóżmy również, że klasy C_1 i C_2 mogą być rozdzielone za pomocą funkcji dyskryminacyjnej, która – w rozważanym przypadku – jest funkcją liniową:

$$d(X) = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n = w_0 + W^TX \qquad (1)$$

Wektor X będzie zaliczony do klasy C_1, jeśli $\mathrm{d}(X) > 0$, albo do klasy C_2, jeśli $d(X) < 0$. Jeśli $d(X) = 0$, to klasyfikacja może być dokonana na podstawie arbitralnej decyzji lub dodatkowych informacji. Równanie $d(X) = 0$ opisuje granice obszarów klasyfikacji. Granica obszarów klasyfikacji jest w tym przypadku $\mathrm{n} - 1$ wymiarową hiperpłaszczyzną:

$$0 = w_0 + w_1x_1 + w_2x_2 + \dots + w_nx_n = w_0 + {}^TX \qquad (2)$$

gdzie $X \in \mathbb{R}^n$ jest zmienną, $W \in \mathbb{R}^n$ jest wektorem współczynników, natomiast w_0 jest wyrazem wolnym. Przypadek dwuwymiarowy (dwóch cech) jest przedstawiony na rysunku 1. Formuła (2) określa hiperpłaszczyznę w przestrzeni $\mathbb{R}^n$ prostopadłą do wektora W i odległą od początku układu współrzędnych o:

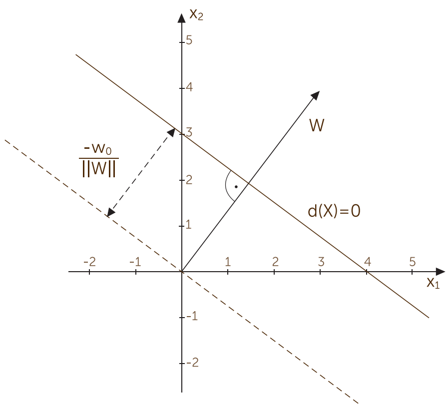
$$\frac{W^TX}{\| W \|} = -\frac{w_0}{\| W \|}$$

gdzie X jest dowolnym punktem hiperpłaszczyzny. Przypadek dwóch cech jest pokazany na rysunku 1. Formułę (1) można przepisać w zmienionej postaci:

$$d(X) = w_0 \cdot 1 + w_1 \cdot x_1 + w_2 \cdot x_2 + \dots + w_n \cdot x_n = \tilde{W}^T\tilde{X}$$

gdzie $X = (1, x_1, x_2, \dots, x_n) \in \mathbb{R}^{n+1}$, $W = [w_0, w_1, w_2, \dots, w_n] \in \mathbb{R}^{n+1}$ są wektorami w $n + 1$ wymiarowej przestrzeni. Dla tej postaci funkcji dyskryminacyjnej równanie granicy obszarów klasyfikacji: $0 = \tilde{W}^T\tilde{X}$

określa n wymiarową hiperpłaszczyznę przechodzącą przez początek układu współrzędnych przestrzeni $\mathbb{R}^{n+1}$.



↑ Rys. 1. Granica $d(X) = 0$ obszarów decyzyjnych równania $d(X) = w_0 \cdot 1 + w_1 \cdot x_1 + w_2 \cdot x_2 = -12 \cdot 1 + 3x_1 + 4x_2$ w układzie współrzędnych OX_1X_2. Wektor wag $W = [3, 4]$ określa orientację prostej rozgraniczającej obszary decyzyjne, wyraz wolny $w_0 = -12$ określa odległość prostej od początku układu współrzędnych

Przypadek wielu klas

W przypadku dwóch klas wystarczyła jedna funkcja liniowa do ich rozdzielenia: elementy były klasyfikowane zgodnie ze znakiem funkcji dyskryminacyjnej. W przypadku wielu klas będziemy stosować jedną funkcję dyskryminacyjną dla jednej klasy:

$$d_i(X) = w_{i0}\cdot 1 + w_{i1}\cdot x_1 + w_{i2}\cdot x_2 + \ldots + w_{ic}\cdot x_c \qquad (3)$$

które powinny spełniać warunki

$$d_i(X) > d_j(X) \text{ dla } j = 1, 2, \ldots, c\ , j \neq i \qquad (4)$$

dla każdego wektora $X \in C_i$. Równanie (3) można przepisać w postaci wektorowej

$$d_i(X) = w_{i0} + W_i^T X \qquad (5)$$

gdzie $W_i = [w_{i1}, w_{i2}, \ldots, w_{in}]$ jest wektorem wag. Granica obszarów decyzyjnych rozdzielająca klasy C_k i C_l będzie oczywiście określona równaniem:

$$(W_k - W_l)^T X + (w_{k0} - w_{l0}) = 0.$$

Przez analogię do przypadku dwóch klas granica ta jest $n-1$ wymiarową hiperpłaszczyzną w przestrzeni $\mathbb{R}^n$. Orientację hiperpłaszczyzny i jej odległość od początku układu współrzędnych określają następujące wielkości:

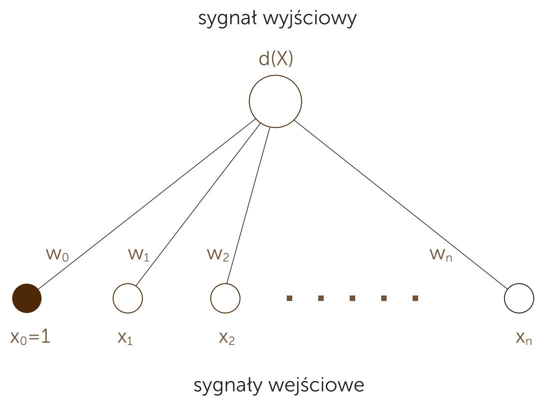
$$W_k - W_l \text{ oraz } \frac{w_{k0} - w_{l0}}{\| W_k - W_l \|}$$

Klasyfikację do wielu klas za pomocą liniowych funkcji decyzyjnych można wyrazić jako sztuczną sieć neuronową – perceptron jednowarstwowy o wielu wyjściach – przedstawioną na rysunku 3. Zakładając, że wagi połączeń zostały określone, obliczenie sieci będzie polegać na wyznaczeniu sygnałów wyjściowych dla każdego wyjściowego neuronu. Wektor wejściowy zostanie zakwalifikowany do klasy odpowiadającej wyjściowemu neuronowi o największej wartości funkcji aktywacji, co – oczywiście – jest zgodne z klasyfikacją z zastosowaniem funkcji decyzyjnych.

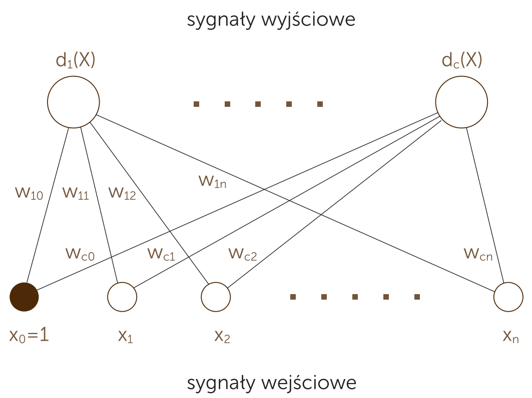
Zauważmy jeszcze, że taka klasyfikacja dzieli przestrzeń wejściową na spójne i wypukłe obszary decyzyjne.

By się o tym przekonać, wystarczy rozważyć dwa punkty X_A i X_B należące do tej samej klasy, tzn. spełniające warunki (4). Punkty X odcinka $\overline{X_A X_B}$ można otrzymać jako kombinację liniową jego końców

$$X = \alpha X_A + (1-\alpha) X_B, \quad 0 \le \alpha \le 1$$



↑ Rys. 2. Reprezentacja liniowej funkcji decyzyjnej jako sztucznej sieci neuronowej. Każdy sygnał wejściowy odpowiada jednej zmiennej funkcji decyzyjnej. Wyraz wolny funkcji decyzyjnej jest reprezentowany waga w_0 i stałym sygnałem wejściowym $x_0 = 1$ oznaczonym kolorem czarnym. Sygnał wyjściowy jest określony wartością funkcji decyzyjnej $d(X) = w_0\cdot 1 + w_1\cdot x_1 + w_2\cdot x_2 + \ldots + w_n\cdot x_n$



↑ Rys. 3. Reprezentacja liniowych funkcji decyzyjnych jako sztucznej sieci neuronowej. Każdy sygnał wejściowy odpowiada jednej zmiennej funkcji decyzyjnej. Sygnał odpowiadający wyrazowi wolnemu jest równy 1 i jest zaznaczony kolorem czarnym. Wartości wyjściowe sieci odpowiadają wartościom poszczególnych funkcji decyzyjnych. Wektor wejściowy jest klasyfikowany do klasy odpowiadającej największej wartości sygnału wyjściowego

Warunki (4) są spełnione dla punktów X odcinka ponieważ są spełnione dla X_A i X_B oraz ponieważ funkcje decyzyjne są funkcjami liniowymi. Dowodzi to spójności i wypukłości obszarów decyzyjnych.

W przypadku dwuwymiarowym z trzema cechami granice obszarów decyzyjnych będą trzema prostymi przecinającymi się w jednym punkcie. Przykładowy podział płaszczyzny jest pokazany na rysunku 4.

Uogólnione funkcje liniowe

Jedną z metod uogólnienia decyzyjnych funkcji liniowych jest przekształcenie wektorów wejściowych za pomocą pewnych r funkcji ϕ zwanych funkcjami bazowymi. Po zastosowaniu funkcji bazowych funkcje decyzyjne przyjmą postać:

$$d_k(X) = \sum_{i=0}^{r} w_{ki}\cdot \phi_i(X),\ k = 1, 2 \ldots, c \qquad (6)$$

gdzie $\phi_i(X)$ jest funkcją bazową i – tej cechy, wyraz wolny w_{k0} jest włączony do wektora wag W. Jeśli funkcje bazowe będą funkcjami wyboru odpowiedniej cechy, to powyższa formuła przyjmie postać zwykłej kombinacji liniowej cech.

Uogólniona aktywacja

W modelu neuronu rozważanym w poprzednim podrozdziale aktywacja neuronu jest funkcją liniową. W wielu modelach neuronów funkcja aktywacji jest złożeniem liniowej funkcji sygnału wejściowego z pewną funkcją monotoniczną:

$$d(X) = g(\tilde{W}^T \tilde{X}) \qquad (7)$$

Liniową funkcję dyskryminacyjną z formuły (1) można przedstawić w postaci uogólnionej przyjmując funkcję g jako funkcję tożsamościową $g(x) = x$.

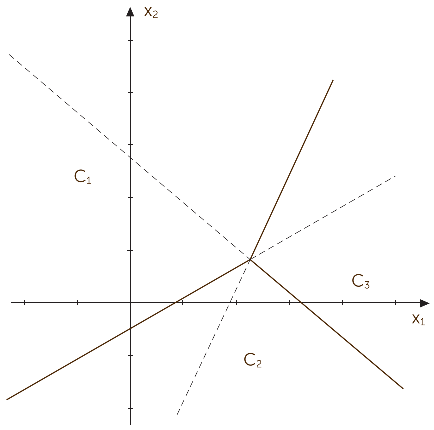
Wprowadzenie monotonicznej modyfikacji sygnału wyjściowego nie zmienia postaci granic obszarów decyzyjnych – nadal będą $n-1$ wymiarowymi hiperpłaszczyznami w przestrzeni $\mathbb{R}^n$.

Aktywacja sigmoidalna. W wielu modelach neuronów aktywacja neuronu jest wynikiem działania funkcji sigmoidalnej unipolarnej na liniową funkcję sygnałów wejściowych:

$$g\colon (-\infty,\infty) \to (0,1),\ g(x) = \frac{1}{1+e^{-x}} \qquad (8)$$

Wersja bipolarna funkcji sigmoidalnej jest określona formułą:

$$g\colon (-\infty,\infty) \to (-1,1),\ g(x) = \frac{1-e^{-x}}{1+e^{-x}} \qquad (9)$$



↑ Rys. 4. Granice obszarów decyzyjnych liniowej sieci jednowarstwowej dla dwóch cech i trzech klas. Proste wyznaczające granice obszarów decyzyjnych są rzutami krawędzi przecięcia płaszczyzn funkcji decyzyjnych na płaszczyznę cech

Ważną własnością funkcji sigmoidalnej jest *standardyzacja* sygnału przez jego homeomorficzne przekształcenie do przedziału jednostkowego. Ponadto, dla niewielkich wartości argumentu, funkcja sigmoidalna jest prawie liniowa, co oznacza, że w takim przypadku neuron jest neuronem liniowym. Wreszcie – funkcje liniowe z sigmoidalną aktywacją są często stosowane w statystyce jako funkcje decyzyjne pod nazwą logicznych funkcji dyskryminacyjnych *logistic discrimination*.

Proces uczenia

Dyskutowany w poprzednich podrozdziałach proces podziału przestrzeni wejściowej na obszary decyzyjne jest mało przydatny w praktyce. By można było zaprojektować sieć neuronową metodami wcześniej dyskutowanymi, to znaczy wyznaczyć wagi połączeń, należy znać obszary decyzyjne. W praktyce zamiast obszarów mamy pewną liczbę wzorców o znanej przynależności do poszczególnych klas. W tej sytuacji należy tak określić poszczególne elementy sieci: architekturę sieci, wagi połączeń, funkcje bazowe, funkcje aktywacji, by zminimalizować błąd klasyfikacji. W niniejszej pracy będziemy dyskutować metody wyznaczania wag połączeń dla arbitralnie określonych architektury sieci, funkcji bazowych i funkcji aktywacji. Zainteresowanych tematyką określania architektury, funkcji bazowych i funkcji aktywacji odsyłamy do bardzo bogatej literatury, np. [1].

Jedną z podstawowych metod wyznaczania przybliżonych wartości wag połączeń sieci jest minimalizacja funkcji błędów klasyfikacji na zbiorze wzorców, zwanym też zbiorem uczącym. Jedną z częściej stosowanych funkcji błędów jest suma kwadratów błędów klasyfikacji na zbiorze uczącym. Sumę tę jako funkcje wag połączeń można wyrazić formułą:

$$E(W) = \tfrac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{c}(d_k(X_i, W) - t_{ki})^2 \qquad (10)$$

gdzie m jest liczbą wszystkich wzorców użytych w procesie uczenia, c – liczba klas, X_i – wektorem cech i-tego wzorca, W – wektorem wag z włączonym wyrazem wolnym w_{k0}, t_{ki} – zadaną wartością funkcji decyzyjnej d_k dla wzorca X_i. Funkcja błędu jest funkcją kwadratową funkcji decyzyjnej. Jeśli funkcja decyzyjna jest gładka, to również funkcja błędu jest gładka. Dla liniowej funkcji decyzyjnej funkcja błędu jest określona formułą:

$$E(W) = \tfrac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{c}\left(\sum_{l=0}^{n} w_{kl}\,\phi_{li} - t_{ki}\right)^2$$

gdzie $\phi_{li} = \phi_l(X_i)$. Różniczkując tę formułę po wagach połączeń i przyrównując pochodne do zera, otrzymamy układ równań dla problemu minimalizacji kwadratów błędów:

$$\sum_{i=1}^{m}\left(\sum_{j=0}^{n} w_{kj}\,\phi_{ji} - t_{ki}\right)\phi_{li} = 0$$

Zapisując to równanie w postaci macierzowej, dostaniemy:

$$(\Phi^T\Phi)\, W^T = \Phi^T\, T \qquad (11)$$

gdzie Φ jest macierzą o wymiarach $n \times m$ z elementami równymi wartościom funkcji decyzyjnych ϕ_{li}, W – macierzą o wymiarach $c \times n$ z elementami w_{ki}, natomiast T – macierzą o wymiarach $m \times c$ z elementami t_{kn}. W równaniu tym macierz Φ z reguły nie jest macierzą kwadratową, dlatego równania nie można rozwiązać przez pomnożenie obu stron przez odwrotność tej macierzy. Natomiast macierz $\Phi^T\Phi$ jest macierzą kwadratową o wymiarach $n \times n$ i jeśli jest macierzą nieosobliwą, to rozwiązaniem równania będzie:

$$W^T = (\Phi^T\Phi)^{-1}\Phi^T\, T$$

gdzie macierz $(\Phi^T\Phi)^{-1}\Phi^T$ jest nazywana pseudoodwrotną do macierzy Φ gdyż, przy założeniu nieosobliwości macierzy $\Phi^T\Phi$, iloczyn macierzy $(\Phi^T\Phi)^{-1}\Phi^T$ i Φ jest macierzą jednostkową. Gdy macierz $\Phi^T\Phi$ jest macierzą osobliwą, rozwiązanie równania jest niejednoznaczne. W praktyce metoda wyznaczania macierzy wag połączeń przez rozwiązanie równań macierzowych jest najczęściej obliczeniowo nieefektywna i nieodporna na błędy numeryczne. Dlatego ta metoda ma znaczenie raczej teoretyczne niż praktyczne. W praktyce do wyznaczenia wag dla danego zbioru wzorców stosuje się metody iteracyjne. W niniejszej pracy przedstawimy gradientową metodę największego spadku.

Niech $E(W)$ oznacza sumę kwadratów błędów klasyfikacji jako funkcje wag połączeń W. Naszym zadaniem jest minimalizacja funkcji błędu E. Zakładając, że funkcja błędu E jest różniczkowalna względem wag, możemy zastosować znaną metodę największego spadku w celu zmniejszenia funkcji błędu. Po pierwsze, inicjujemy arbitralnie lub losowo wartości wag. Oznaczmy wartości początkowe wag przez $W^{(0)}$, gdzie górny indeks oznacza kolejne przybliżenie wartości wag. Dla danego przybliżenia $W^{(r)}$ wartości wag wyznaczamy gradient funkcji błędu $\nabla_W E$ w przestrzeni wag, następnie modyfikujemy wektor wag:

$$W^{(r+1)} = W^{(r)} - \eta\, \nabla_W E(W^{(r)}) \qquad (12)$$

gdzie η jest *parametrem uczenia*. Zmiana wektora wag polega na niewielkim przesunięciu w kierunku przeciwnym do wskazywanego przez gradient. Ponieważ gradient wskazuje kierunek najszybszego wzrostu funkcji, więc taka modyfikacja zmniejszy wartość funkcji błędu w kierunku największego spadku. Powtarzając ten proces wielokrotnie, otrzymamy ciąg wag $W^{(r)}$ oraz ciąg wartości błędu klasyfikacji $E(W^{(r)})$. Dla odpowiedniej wartości parametru uczenia ciąg wartości funkcji błędu powinien być ciągiem malejącym i szybko zmierzać do minimum funkcji błędu. Istotną rolę pełni dobór wartości parametru uczenia. Dla zbyt małej wartości parametru uczenia ciąg wartości funkcji błędu będzie ciągiem malejącym, ale bardzo wolno zbieżnym. Natomiast dla zbyt dużej wartości parametru uczenia zamiast zbieżności pojawią się oscylacje wokół minimum funkcji błędu.

W procesie uczenia będziemy korzystać z reguły uczenia dla pojedynczych wag, zamiast z reguły wektorowej. Wektorowa reguła uczenia (12) jest rozpisywana na poszczególne wagi następująco:

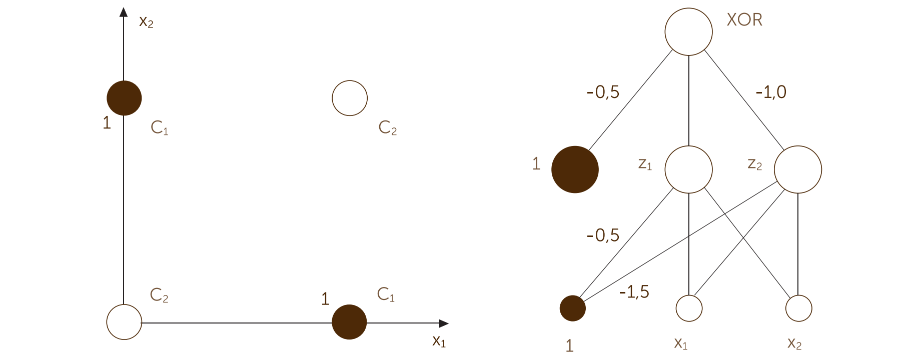
$$w_{kj}^{(r+1)} = w_{kj}^{(r)} - \eta\,\frac{\partial E}{\partial w_{kj}}(W^{(r)})$$

Formuła (13), określająca łączny błąd klasyfikacji dla wszystkich wzorców, jest sumą błędów klasyfikacji $E^{(i)}$ dla poszczególnych wzorców, gdzie:

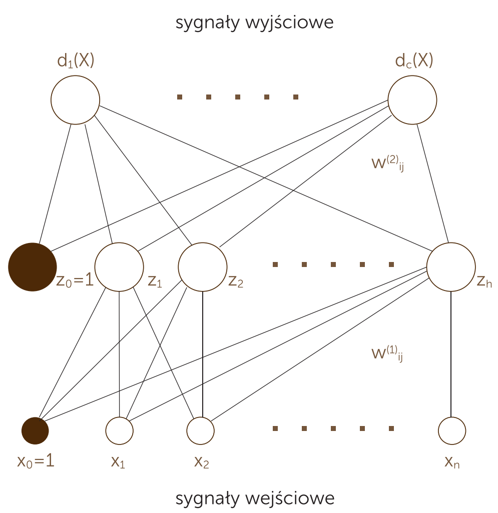
$$E^{(i)}(W) = \tfrac{1}{2}\sum_{k=1}^{c}(d_k(X_i, W) - t_{ki})^2,\quad i = 1, 2, \ldots, m \qquad (13)$$

Upraszczając proces uczenia, będziemy modyfikować wagi dla poszczególnych wzorców:

→



↑ Rys. 5. Funkcja $\mathrm{XOR}(x_1, x_2)$. Klasa C_1 zawiera punkty $(0, 1)$ i $(1, 0)$, dla których funkcja przyjmuje wartość 1. Klasa C_2 zawiera punkty $(0, 0)$ i $(1, 1)$, dla których funkcja przyjmuje wartość 0. Spójne i wypukłe obszary decyzyjne liniowej sieci nie rozdziela funkcji XOR. Dwuwarstwowa sieć liniowa przedstawiona z prawej strony realizuje funkcje XOR. Wagi połączeń są równe 1 albo są określone na rysunku. Funkcja aktywacji neuronów warstw ukrytej i wyjściowej jest kombinacją liniową sygnałów wejściowych zmodyfikowaną unipolarną funkcją progową określoną formułą (0.17)



↑ Rys. 6. Dwuwarstwowa sieć neuronowa. Sygnały wejściowe x_i są podawane na wejście neuronów warstwy ukrytej. Wyniki z_i obliczenia neuronów warstwy ukrytej są następnie podawane na wejście neuronów warstwy wyjściowej

$$w_{kj}^{(i)(r+1)} = w_{kj}^{(i-1)(r+1)} - \eta \frac{\partial E^{(i)}}{\partial w_{kj}} W^{(i-1)(r+1)},$$

$$i = 1, 2, \dots, m \tag{14}$$

gdzie: $w_{kj}^{(0)(r+1)}$, $W^{(0)(r+1)}$ są wagami otrzymanymi po r cyklach modyfikacji, $w_{kj}^{(i-1)(r+1)}$, $W^{(i-1)(r+1)}$ – wagami otrzymanymi po i wzorcach w $r+1$ cyklu modyfikacji, $w_{kj}^{(m)(r+1)}$, $W^{(i-1)(r+1)}$ – wagami po $r+1$ cyklach modyfikacji.

W tym przypadku modyfikacje wag będą dokonywane kolejno dla wszystkich wzorców. Proces ten będzie powtarzany cyklicznie aż do zakończenia procesu uczenia sieci.

Rozważmy jeszcze pochodną funkcji błędu reguły uczenia biorąc pod uwagę uogólnioną funkcję aktywacji (7). Założenie różniczkowalności funkcji błędu wymaga również różniczkowalności funkcji g modyfikującej kombinację liniową sygnałów wejściowych.

Załóżmy chwilowo, że funkcja aktywacji jest kombinacją liniową sygnałów wejściowych, to znaczy funkcją g w równaniu (7) jest funkcja identycznościowa. Pochodna funkcji błędu dla wzorca X_i będzie wówczas wyrażona formułą:

$$\frac{\partial E^{(i)}}{\partial w_{kj}} = (d_k(X_i, W) - t_{ki})\, \phi_j(X_n) = \delta_{ki} \cdot \phi_{ji} \tag{15}$$

gdzie $\phi_j(X_n)$ jest funkcją bazową ϕ_j zastosowaną do wektora X_i, $\delta_{ki} = d_k(X_i, W) - t_{ki}$, $\phi_{ji} = \phi_j(X_n)$.

Natomiast reguła modyfikacji wag (14) przyjmie postać:

$$w_{kj}^{(r+1)} = w_{kj}^{(r)} - \eta \delta_{ki}\, \phi_{ji}$$

Jeśli funkcja g jest funkcją różniczkowalną, to w powyższych wzorach należy uwzględnić jej wpływ na pochodne funkcji błędu:

$$\frac{\partial E^{(i)}}{\partial w_{kj}} = g'(a_k) \delta_{ki}\, \phi_{ji} \tag{16}$$

gdzie

$$a_k = \sum_{j=0}^{n} w_{kj} \phi_j\,, \quad \delta_{ki} = (d_k(X_i) - t_{ki})$$

Sieć wielowarstwowa – rozdzielenie bez ograniczeń

Liniowa sieć jednowarstwowa omawiana w poprzednim rozdziale ma niewielkie zastosowanie praktyczne. Rzeczywiste zagadnienia klasyfikacji, które mogą być rozwiązane za pomocą takiej sieci są raczej wyjątkami. Warunek wypukłości i spójności obszarów decyzyjnych jest bardzo mocno ograniczający. Przykładowo, sieć jednowarstwowa nie może implementować funkcji logicznej XOR, porównaj rysunek 5. W tym rozdziale zajmiemy się sieciami wielowarstwowymi, w których neurony będą podzielone na kolejne warstwy. Założymy, że w rozważanych sieciach wielowarstwowych istnieją połączenia tylko między sąsiednimi warstwami. W sieciach wielowarstwowych ostatnia warstwa neuronów nazywa się warstwą wyjściową, natomiast poprzednie warstwy nazywają się warstwami ukrytymi. W sieci jednowarstwowej nie ma warstw ukrytych. Na rysunku 6 przedstawiono sieć dwuwarstwową. Sygnały wejściowe x_i są podawane na wejście neuronów warstwy ukrytej.

Sygnały wyjściowe z_i neuronów warstwy ukrytej są podawane na wejście neuronów warstwy wyjściowej.

$$g(x) = \begin{cases} 1 & \text{dla} \quad x \geqslant 0 \\ 0 & \text{dla} \quad x < 0 \end{cases} \tag{17}$$

Sieć dwuwarstwowa

Załóżmy, że mamy sieć dwuwarstwową jak na rysunku 6. Wartość sygnałów wyjściowych neuronów warstwy ukrytej jest kombinacją liniową sygnałów wejściowych opisaną formułą:

$$a_i = \sum_{j=0}^{n} w_{ij}^{(1)} x_j\,, \; i = 1, 2, \dots, h \tag{18}$$

gdzie sygnał stały został włączony do sumy z wartością równą 1. Sygnały wyjściowe neuronów warstwy ukrytej są następnie poddawane działaniu funkcji aktywacji, dając:

$$z_i = g(a_i) = g\,(\sum_{j=0}^{n} w_{ij}^{(1)} x_j),$$

$$i = 1, 2, \dots, h \tag{19}$$

Wartości aktywacji funkcji są następnie przekazywane jako sygnały wejściowe neuronów warstwy wyjściowej. Sygnały wyjściowe neuronów warstwy wyjściowej są, jak poprzednio, kombinacją liniową sygnałów wejściowych tych neuronów:

$$b_i = \sum_{j=0}^{h} w_{ij}^{(2)} x_j, \;\; i = 1, 2, \dots, c \tag{20}$$

gdzie, tak jak poprzednio, sygnał stały został włączony do sumy z wartością równą 1.

Ostatecznie, sygnały wyjściowe całej sieci są wynikiem działania funkcji aktywacji neuronów warstwy wyjściowej na ich sygnały wyjściowe:

$$d_k = \tilde{g}(b_k) = \tilde{g}\ (\sum_{j=0}^{h} w_{kj}^{(2)} x_j),$$
$$k = 1, 2, \ldots, c \qquad (21)$$

Łącząc powyższe formuły, otrzymamy następujący opis sieci:

$$d_k = \tilde{g}\ (\sum_{l=0}^{h} w_{kl}^{(2)}\, g\ (\sum_{j=0}^{n} w_{lj}^{(1)}\ x_j)),$$
$$k = 1, 2, \ldots, c \qquad (22)$$

Zasadniczą jednak różnicą miedzy siecią jednowarstwową i wielowarstwową jest możliwość adaptacyjnego ustalania wag połączeń poszczególnych warstw. W sieci jednowarstwowej i sieciach wielowarstwowych równoważnych sieciom jednowarstwowym adaptacja wag polaczeń jest efektywnie dokonywana tylko w jednej warstwie. Zanim jednak rozważymy problem adaptacji wag połączeń, to znaczy – problem uczenia sieci, zastanowimy się nad możliwościami decyzyjnymi sieci wielowarstwowych.

Sieć wielowarstwowa

Przypomnijmy, że sieć jednowarstwowa dla przypadku dwóch klas dzieliła przestrzeń $\mathbb{R}^n$ sygnałów wejściowych na dwie podprzestrzenie rozdzielone hiperpłaszczyzną $n - 1$ wymiarową. Przypadek dwuwymiarowej przestrzeni sygnałów wejściowych jest pokazany na rysunku 7a.

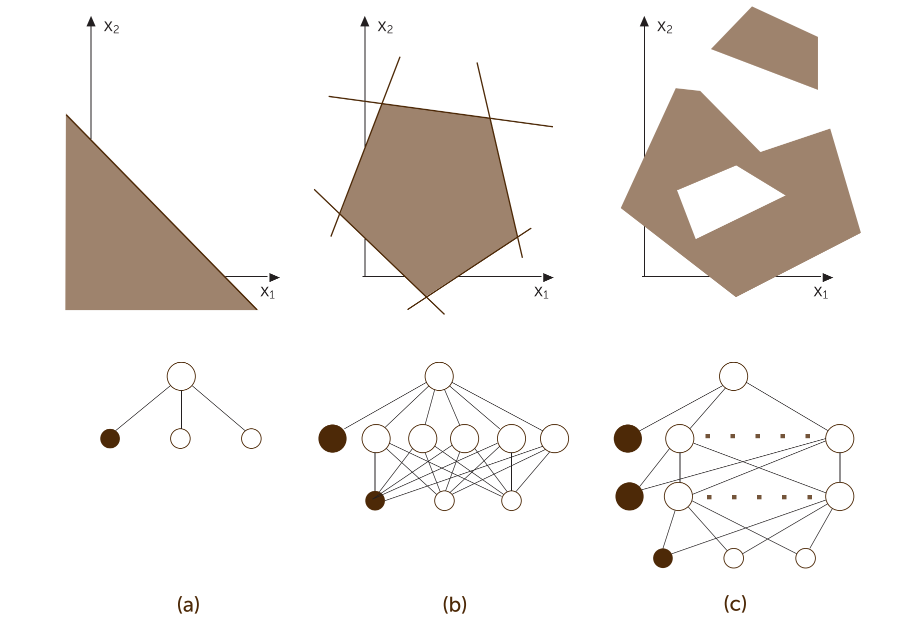
Sieć jednowarstwowa jest siecią ograniczoną do wypukłych obszarów decyzyjnych. Dlatego na przykład sieć jednowarstwowa nie może realizować dowolnych funkcji logicznych. Natomiast zastosowanie sieci dwuwarstwowej z funkcją aktywacji neuronów warstw ukrytej i wyjściowej umożliwia realizację dowolnej funkcji logicznej, porównaj [8]. Konstrukcja takiej sieci jest prosta. Zauważmy, że w przypadku n binarnych (logicznych) sygnałów wejściowych, wszystkich możliwych wartości sygnałów wejściowych będzie 2^n, nazwijmy je wzorcami. Żeby określić funkcje logiczną dla n zmiennych, wystarczy podać wartość (binarną, logiczną) tej funkcji dla dowolnego wzorca (podać wartości jej zmiennych). Konstruując sieć obliczającą funkcję logiczną, należy umieścić $k+1$ neuronów w warstwie ukrytej, gdzie k jest liczbą wzorców, dla których funkcja przyjmuje wartość 1. Dodatkowy neuron jest neuronem sygnału stałego. Kombinacja liniowa sygnałów wejściowych jest modyfikowaną unipolarną progową funkcją aktywacji określoną równaniem (17). Rozważmy pewien wzorzec, dla którego wartość obliczanej funkcji jest równa 1. Wagi połączeń sygnałów wejściowych z odpowiadającym mu neuronem warstwy ukrytej przyjmujemy równe 1 dla sygnałów równych 1 oraz przyjmujemy równe -1 dla sygnałów równych 0. Określając wartość wagi sygnału stałego równą $-t$, gdzie t jest liczbą sygnałów równych 1 w tym wzorcu, otrzymamy neuron generujący wartość 1 tylko dla tego wzorca. Zatem, dokładnie jeden z neuronów warstwy ukrytej wygeneruje sygnał równy 1 wtedy i tylko wtedy, gdy na wejściu sieci pojawi się wzorzec, dla którego wartość obliczanej funkcji jest równa 1. W pozostałych przypadkach wszystkie neurony warstwy ukrytej będą generować sygnał 0. Wystarczy teraz przyjąć wagi połączeń neuronów warstwy ukrytej z neuronem wyjściowym równe 1 oraz wagę sygnału stałego równą -1, by neuron wyjściowy z unipolarną progową funkcją aktywacji generował sygnał równy 1 wtedy i tylko wtedy, gdy wartość obliczanej funkcji jest równa 1.

Dwuklasowa sieć dwuwarstwowa pozwala na konstruowanie obszarów decyzyjnych o bardziej złożonych kształtach, porównaj [5, 6]. Jeśli przyjmiemy progową unipolarną funkcję aktywacji neuronów warstwy ukrytej i neuronu wyjściowego, to dla wartości wagi $w_{01}^{2} = -h$, gdzie h jest liczbą neuronów warstwy ukrytej, neuron wyjściowy będzie obliczał funkcję *and*. Przyjmując odpowiednie skierowanie obszarów decyzyjnych neuronów warstwy ukrytej, otrzymamy obszar decyzyjny w postaci dowolnego jednospójnego wieloboku. Na rysunku 7b jest przedstawiona realizacja obszaru decyzyjnego w postaci wypukłego pięciokąta.

Sieć trzywarstwowa z unipolarną progową funkcją aktywacji może realizować dowolne obszary decyzyjne z dowolnie dużą dokładnością, porównaj [5]. Dwie pierwsze warstwy realizują jednospójne składniki obszarów decyzyjnych, natomiast warstwa trzecia realizuje włączanie–wyłączanie składników, tak by otrzymać przybliżenie obszaru decyzyjnego. Dokładność przybliżenia jest kwestią dokładności aproksymacji jednospójnych składowych za pomocą wieloboków. Rysunek 7c ilustruje możliwości decyzyjne sieci trzywarstwowej. Obszar decyzyjny składa się z czworoboku oraz siedmioboku w wyłączonym czworobokiem.

Powyższe rozważania przybliżają idee dowodu stwierdzenia, ze trzywarstwowa sieć może realizować dowolną funkcję ciągłą sygnałów wejściowych określoną na ograniczonej i domkniętej dziedzinie. Dowód tego stwierdzenia wynika z twierdzenia Kołmogorowa (przeczącego trzynastemu problemowi Hilberta) opublikowanego w [4]. Wersje dowodu Kołmogorowa, które można interpretować na gruncie sieci neuronowych z ciągłą funkcją aktywacji jako dowody istnienia sieci realizujących funkcje decyzyjne zostały podane w pracach [7, 14].

Powyższe rozważania nie tylko dowodzą istnienia sieci realizujących funkcje



↑ Rysunek 7. Ilustracja możliwości decyzyjnych sieci jednowarstwowej (a), dwuwarstwowej (b) i trzywarstwowej (c) decyzyjne, ale też pozwalają wyznaczyć parametry – wagi polaczeń – takich sieci. Dowody te w sposób istotny umożliwiły również zrozumienie działania sztucznych sieci neuronowych. Jednak ich znaczenie praktyczne jest niewielkie, szczególnie gdy staniemy przed zadaniem konstrukcji sieci przy zadanym zbiorze wzorców.

Proces uczenia

Podobnie jak w przypadku sieci jednowarstwowej proces uczenia będzie oparty na minimalizacji funkcji błędów klasyfikacji na zbiorze uczącym. Minimalizacja funkcji będzie dokonywana na względem wartości wag połączeń. W procesie uczenia sieci wielowarstwowej należy zwrócić uwagę na funkcję aktywacji. Jeśli proces uczenia jest oparty na metodzie gradientowej, a tak jest w znakomitej większości przypadków, to funkcje aktywacji powinny być funkcjami różniczkowalnymi. Co prawda, w przypadku progowej funkcji aktywacji, do ostatniej warstwy połączeń można zastosować perceptronową regułę uczenia, ale nie można jej przenieść do warstw poprzednich, gdyż nie mamy metody określania wpływu sygnałów pochodzących z neuronów poprzedniej warstwy – przy progowej funkcji aktywacji – na błędy generowane na wyjściu sieci. Dlatego, chcąc stosować metody gradientowe, powinniśmy zakładać różniczkowalność funkcji aktywacji. Przy tych założeniach można będzie wyznaczyć pochodne funkcji błędu względem wag połączeń poszczególnych warstw, a następnie zastosować metodę największego spadku lub bardziej wyrafinowane metody optymalizacyjne do zmniejszenia wartości funkcji błędu.

Jako funkcję błędu zastosujemy sumę kwadratów błędów określoną formułą (13). W celu uproszczenia rozważań założymy, że sygnał wyjściowy każdego neuronu jest opisany formułą (7) z różniczkowalną funkcją aktywacji g. Proces uczenia będzie rozumiany jako korekty wag połączeń dla każdego sygnału wyjściowego i dla każdego wzorca ze zbioru uczącego. By uprościć zapis załóżmy, że sieć jest siecią dwuwarstwową. Biorąc pod uwagę formułę (21), możemy łatwo wyznaczyć pochodną sygnału wyjściowego względem wag drugiej warstwy przez analogie do formuły (16):

$$\frac{\partial E^{(i)}}{\partial w_{kj}^{(2)}} = \delta_{ki}\, \tilde{g}'(b_k) z_j \qquad (23)$$

gdzie:

$$\delta_{ki} = (d_k(X_i) - t_{ki}),\ b_k = \sum_{l=0}^{h} w_{kl}^{(2)} z_l$$

z_j jest sygnałem wejściowym neuronu o numerze j (warstwy wyjściowej), a pozostałe oznaczenia są zgodne z przyjętymi w formule (16). Oznaczając $\delta_{ki}\, \tilde{g}'(b_k)$ przez $\lambda_k^{(1)}$, otrzymamy uproszczoną postać formuły :

$$\frac{\partial E^{(i)}}{\partial w_{kj}^{(2)}} = \lambda_k^{(1)} z_j \qquad (24)$$

Podobnie, bazując na formule (22), można wyznaczyć pochodne funkcji błędu względem wag połączeń pierwszej warstwy:

$$\frac{\partial E^{(i)}}{\partial w_{lj}^{(1)}} = \sum_{k=0}^{c} \delta_{ki}\, \tilde{g}'(b_k) \frac{\partial b_k}{\partial w_{lj}^{(1)}} \qquad (25)$$

gdzie:

$$b_k = \sum_{p=0}^{h} w_{kp}^{(2)} z_p = \sum_{p=0}^{h} w_{kp}^{(2)} g\,(\sum_{q=0}^{n} w_{pq}^{(1)} g\, x_q)$$

Stąd:

$$\frac{\partial E^{(i)}}{\partial w_{lj}^{(1)}} = (\sum_{k=0}^{c} \delta_{ki}\, \tilde{g}'(b_k) w_{kl}^{(2)}\, g'(a_l)) x_j \qquad (26)$$

gdzie:

$$\delta_{ki} = \ (d_k(X_i) - t_{ki}),\ \ b_k = \sum_{p=0}^{h} w_{kp}^{(2)} z_p,$$

$$a_l = \sum_{r=0}^{n} w_{lr}^{(1)} x_r$$

Przekształcając formułę (26) i stosując oznaczenie $\lambda_k^{(2)} = \delta_{ki}\, \tilde{g}'(b_k)$, otrzymamy:

$$\frac{\partial E^{(i)}}{\partial w_{lj}^{(1)}} = g'(a_l) \sum_{k=0}^{c} w_{kl}^{(2)} \lambda_k^{(2)} x_j \qquad (27)$$

Oznaczając:

$$g'(a_l)(\sum_{k=0}^{c} w_{kl}^{(2)} \lambda_k^{(2)}) = \lambda_l^{(1)}$$

otrzymujemy formułę identyczną z formułą (24):

$$\frac{\partial E^{(i)}}{\partial w_{lj}^{(1)}} = \lambda_l^{(1)} x_j$$

Zauważmy, że $\lambda_l^{(1)}$ jest kombinacją liniową współczynników $\lambda_k^{(2)}$ wyznaczonych dla wszystkich neuronów następnej warstwy (bliższej warstwie wyjściowej, w naszym przypadku warstwy wyjściowej). Uogólnienie tych formuł na sieci wielowarstwowe jest natychmiastowe. W ten sposób otrzymujemy prostą metodę wyznaczania pochodnych funkcji błędu względem wag połączeń, a zatem możemy zastosować gradientowe metody minimalizacji funkcji błędu.

Powyższa dyskusja metod uczenia sieci neuronowych oczywiście nie wyczerpuje bardzo obszernej dziedziny i nie podaje wielu interesujących szczegółów. Informacje na ten temat można znaleźć w bogatej literaturze, na przykład w [1].

LITERATURA

[1]. Bishop C. M., *Neural Networks for Pattern recognition*, Oxford University Press, 1997.

[2]. Hebb D. O., *The Organization of Behavior: A Neuropsychological Theory*, John Wiley & Sons, New York, 1949.

[3]. Homenda W., *Zastosowanie inteligentnych metod obliczeniowych do przetwarzania informacji muzycznych*, Akademicka Oficyna Wydawnicza EXIT, seria: Badania systemowe, ISBN 83-60434-28-4, tom 56, 1-260, 2007.

[4]. Kołmogorow A. N., *On the representation of continuous functions of several variables by superposition of continuous functions of one variable and addition*, Doklady Akademii Nauk SSSR, 114(5)1957, 953-956.

[5]. Lippmann R. P., *An introduction to computing with neural nets*, IEEE ASSP Magazine, April, 1987, 4-22.

[6]. Lonstaff I. D. and Cross J. F., A pattern recognition approach to understanding the multi-layer perceptron, Pattern Recognition Letters, vol. 5, 1987, 315-319.

[7]. Lorentz G. G., *On the 13th problem of Hilbert*, in: Proceedings of Symposia in Pure Mathematics, pp. 419-429, Providence, RI: American Mathematical Society, 1976.

[8]. McCulloch W. S. and Pitts W. H., *A Logical Calculus of the Ideas Imminent in Nervous Activity*, Bulletin oh Mathematical Biophysics, 5, 1943, 115-133.

[9]. Minsky M. and Papert S., Perceptron: *An Introduction to Computational Geometry*, The MIT Press, Cambridge, Massachusetts, 1969.

[10]. Rosenblatt E., *Two Theorems of Statistical Separability in the Perceptron, in: Mechanization of Thought Process*: Proc. of Symposium No. 10, pp. 421-456, National Physical Laboratory, London, November 1958.

[11]. Rosenblatt E., *Principles of Neurodynamics: Perceptrons and the theory of Brain Mechanisms*, Spartan, Washington, D.C., 1962.

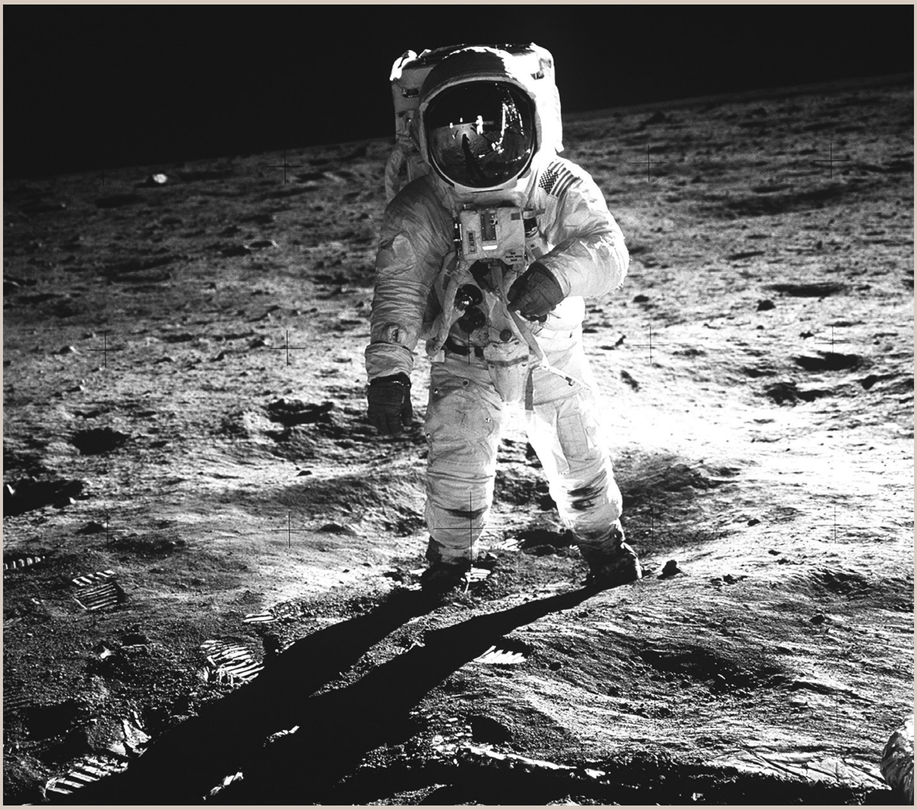
[12]. Rumelhart D. E., Hinton G. E. and Williams R.J., *Learning Internal Representations by Error Propagation, in: Parallel Distributed Processing*: Explorations in the Microstructures of Cognition, vol 1: Foundations, Rumelhart D. E. et al. (Eds.), The MIT Press, Cambridge, Massachusetts, 1986.

[13]. Simon J.C., *Patterns and Operators: The Foundations of Data Representations*, McGraw-Hill, New York, 1986.

[14]. Sprecher D. A., *On the structure of continuous functions of several variables*, Transactions of the American Mathematical Society, 115, 1965, 340-355.

{**Profesor Władysław Homenda** pracuje na Wydziale Matematyki i Nauk Informacyjnych Politechniki Warszawskiej, kierownik Zakładu Strukturalnych Metod Przetwarzania Wiedzy. Zajmuje się m.in. złożonością obliczeniową, algorytmami, teorią automatów, granicami obliczalności, inteligentnymi technologiami obliczeniowymi. Od wielu lat współpracuje z Centrum Studiów Zaawansowanych jako wykładowca Uczelnianej Oferty Dydaktycznej.}

PRZEKRACZANIE GRANIC



↑ Astronauta Buzz Aldrin na powierzchni Księżyca. źródło: NASA - http://grin.hq.nasa.gov/ABSTRACTS/GPN-2001-000013.html, Domena publiczna, https://commons.wikimedia.org/w/index.php?curid=32049

„...jesteśmy stworzeniami ciekawskimi, dociekliwymi, kreatywnymi, o niezaspokojonym głodzie wiedzy, przesuwającymi granice poznania..."

Pytanie o sens życia wydaje się być jednocześnie najważniejszym i jednocześnie najbardziej pomijanym w egzystencji człowieka. W naszej naturze leży przecież dociekanie, badanie i szukanie celu, a jednak wielu ludzi zdaje się nie przywiązywać wagi do tego poszukiwania. Pierwszą odpowiedzią, która nasuwa mi się, gdy ktoś zapyta mnie o sens życia, jest niejako wyuczona, czy może raczej wrośnięta we mnie odpowiedź wynikająca z mojej wiary: „osiągnięcie zbawienia i życia wiecznego". Mówię „wrośnięta", gdyż mimo, że szczerze wierzę, iż przestrzeganie przykazań Bożych i nauki Chrystusa powinno wyznaczać kierunek ludzkiemu istnieniu, to ciężko jest nam – ludziom młodym, studentom – przyjąć to jako nadrzędny cel. Łatwo ulegamy pokusom świata materialnego, w tym wyznaczaniu sobie materialnych punktów docelowych. Dążenie do bycia najlepszym, do posiadania, do zajmowania jakiejś funkcji – to wszystko ambicje materialne (nie duchowe). Czym więc dla mnie jest sens życia w ujęciu „ziemskim"?

Niedawno zacząłem interesować się nieco podróżami kosmicznymi. Znajomość ich historii wydaje mi się niesamowicie niedouczaną i pomijaną gałęzią edukacji, jednocześnie bardzo edukującą i motywującą. Niesamowite wrażenie zrobiło na mnie uświadomienie sobie, że niewiele ponad 60 lat po tym, jak pierwszy samolot oderwał się od ziemi, człowiek postawił nogę na księżycu! Wydaje mi się, że był to jeden z tych momentów (zdanie sobie sprawy, jak niewiele czasu dzieliło te wydarzenia), które zmieniają życie. Uświadomiłem sobie wtedy, kim jest człowiek: jesteśmy stworzeniami ciekawskimi, dociekliwymi, kreatywnymi, o niezaspokojonym głodzie wiedzy, przesuwającymi granice poznania. Z naszej istoty i dotychczasowych dokonań wynika więc zobowiązanie. Powinniśmy zdać sobie sprawę, jaka jest nasza rola w społeczeństwie, jak możemy pomóc ludzkości, a następnie robić wszystko, aby tę rolę wypełniać jak najlepiej.

Jestem świadomy, że myśli te brzmią mocno utopijnie, ale czy bez dążenia do stworzenia idealnego świata, można takowy stworzyć? Czy nie mając wizji da się przypadkiem dojść do czegoś wielkiego? Zgodnie ze Starym Testamentem jesteśmy stworzeni na obraz i podobieństwo Boga, naszym zadaniem zdaje się więc być dążenie do bycia Mu podobnymi. Oznacza to pracę nad naszym ludzkim pojmowaniem dobroci, miłości, doskonałości, sprawiedliwości, ale także (co może zabrzmieć nieco absurdalnie) przekraczanie granic czasu, przestrzeni i innych nieodkrytych jeszcze przez nas wymiarów. Jestem pewien, że żadne z tych celów nie zostaną osiągnięte za mojego życia, niemniej jednak czekam z niecierpliwością, aby zobaczyć (być może z zaświatów) w jakim kierunku rozwinie się ludzkość i jaką przyszłość sobie zgotujemy.

{Przemysław Szczeciński - Wydział Matematyki i Nauk Informacyjnych PW}

LĘK A WOLNOŚĆ

W ostatnim czasie wiele myślałam o lęku, o tym co go powoduje i jak z nim walczyć. Nie mam tu na myśli stresu tuż przed egzaminem, a lęk, który pojawia się gdzieś głęboko w nas i który blokuje nas przed podejmowaniem decyzji i nowych wyzwań. Jednym słowem nie pozwala nam żyć w pełni. Nie można być prawdziwie wolnym człowiekiem i jednocześnie żyć w lęku. Strach często spowodowany jest przez schematy i stereotypy którymi jesteśmy otoczeni na co dzień. Tak naprawdę każdego dnia musimy walczyć, aby właśnie różnorodne schematy obalać i stawać się naprawdę wolnym człowiekiem. Duży problem stanowią wypowiedzi ludzi, którzy nie mając dostatecznej wiedzy na dany temat głoszą fałszywe tezy. Wprowadzają w ten sposób inne osoby w błąd, zniekształcają obraz rzeczywistości, a przez to budzą lęk. Przykładem mogą być próby umoralniania innych, mówienia im, że to co robią jest złe, chociaż tak naprawdę złem to nie jest. Takie mity zazwyczaj dotyczą delikatnej sfery życia duchowego. Jednakże na innych płaszczyznach życia również można spotkać się z zakłamaniem rzeczywistości. Po pierwszym roku na studiach magisterskich na MiNI jestem bardzo pozytywnie zaskoczona, jak wiele rzeczy o których kiedyś słyszałam i których się bałam, okazało się być zupełnym kłamstwem. Takie mity mogą być jednak bardzo niszczące, bo często odstraszają innych ludzi od czegoś, co akurat dla nich było by dobre.

Inną przyczyną społecznego lęku są właśnie stereotypy. Jeśli kogoś nie znamy, to najłatwiej umieścić go w szufladce z grupą ludzi, do której najbardziej według nas pasuje. Tak powstaje chociażby lęk przed imigrantami, bezdomnymi czy po prostu innymi kulturami. Wyobrażenia o drugim człowieku blokują nas przed prawdziwym spotkaniem z nim, a co za tym idzie przed poznaniem jego faktycznej natury. Od dziecka zostają nam również wpojone różne schematy tego jak dobrze żyć, jak się zachowywać. Dają nam one złudne poczucie bezpieczeństwa. Może nam się wydawać, że podążając utartą drogą, będziemy mieć spokojne i szczęśliwe życie. Niestety później życie wszystko weryfikuje i zazwyczaj okazuje się, że te schematy zupełnie nie działają i żeby żyć naprawdę trzeba szukać drogi jak się od nich uwolnić.

↑ Leszek Sokół, obraz zatytułowany *Wolność kocham i rozumiem*

„… bycie pewnym siebie oznacza, że mamy większą szansę na sukces"

Jeżeli żyjemy zgodnie z jakimś schematem, bez własnej inwencji, to wtedy często pojawia się lęk, że na tym świecie nie ma dla nas miejsca. Obawa, że nie możemy tu na Ziemi robić tego co kochamy. Lęk może pojawiać się również, gdy nie czujemy się wystarczająco dobrzy, mądrzy, kochani. Wtedy zazwyczaj zamykamy się w sobie. Unikamy kontaktu z ludźmi, którzy wydają się być od nas lepsi. Jeżeli nie akceptujemy w pełni siebie, trudno też polubić innych. W wyniku niskiej samooceny lęk paraliżuje nas przed podejmowanie nowych działań. Pozostajemy wtedy przeciętni. A przecież człowiek zobowiązany jest do pracy nad sobą i do mądrego rozwoju.

Strach ma jednak też dobre strony. Modne ostatnio hasło porzucenia własnej strefy komfortu nie zawsze działa. Przecież człowiek nie musi być najlepszy we wszystkim. Ciągłe przekraczanie własnych granic i udowadnianie innym, że jesteśmy panami własnego losu do niczego mądrego moim zdaniem nie prowadzi. Czasami lęk jest ostrzeżeniem, żeby nie iść w jakimś kierunku. Pokazuje nam, że coś innego jest dla nas dobre. Nie zmienia to jednak faktu, że aby czynić w życiu piękne rzeczy potrzebujemy odwagi i pokonania pewnego rodzaju lęku. Wszystkim i sobie życzę tego, abyśmy nie lękali się przyszłości, aby nie towarzyszyły nam obawy, że nasze wartości będą w tym świecie niezrozumiałe, lecz abyśmy pozostali im wierni i walczyli o piękny i mądry świat.

{Karolina Pawlak - Wydział Matematyki i Nauk Informacyjnych Politechniki Warszawskiej}

WINTER WORKSHOP ON COMPLEX SYSTEMS

Pierwsze polskie Zimowe Warsztaty Układów Złożonych za nami!

W dniach 4-8 lutego 2019 w Zakopanem i Krakowie odbyły się, pierwszy raz w Polsce, Zimowe Warsztaty Układów Złożonych (Winter Workshop on Complex Systems http://wwcs2019.org/) organizowane przez Wydział Fizyki Uniwersytetu Warszawskiego, przy wsparciu m.in. Wydziału Fizyki Politechniki Warszawskiej i Centrum Studiów Zaawansowanych PW.

Ideą warsztatów jest stworzenie cyklicznego wydarzenia, które pozwoli na spotkanie młodych naukowców z całego świata. Najważniejszym ich elementem są wspólne projekty naukowe, nad realizacją których czuwają zaproszeni goście – wybitni naukowcy zajmujący się badaniami układów złożonych. W tegorocznej edycji szkoły gośćmi byli: Chiara Poletto (French National Institute of Health and Medical Research, http://chiara-poletto.weebly.com/), Pere Colet (Institute for Cross-Disciplinary Physics and Complex Systems IFISC, Institute for Cross-Disciplinary Physics and Complex Systems IFISC, https://ifisc.uib-csic.es/en/people/pere-colet/), Piotr Fronczak (Politechnika Warszawska, http://www.if.pw.edu.pl/~agatka/index_en.html) i Fabrizio Lillo (University of Bologna, https://fabriziolillo.wordpress.com/). Poza zaproszonymi gośćmi w warsztatach udział wzięło 45 młodych naukowców, głównie doktorantów i osób na stażach podoktorskich, z ponad 15 krajów Europy i świata. Warsztaty, jak cała nauka o złożoności (ang. Complex Systems Science) słyną ze swojego interdyscyplinarnego charakteru, nie inaczej było też w polskiej edycji - gościliśmy koleżanki i kolegów zajmujących się biologią, chemią, ekonomią, fizyką, informatyką, matematyką, medycyną, naukami społecznymi, psychologią czy statystyką, a przeważnie subdyscyplinami na styku kilku wymienionych.

Dzięki wsparciu Centrum Studiów Zaawansowanych mogliśmy nagrodzić najbardziej aktywnych uczestników warsztatów, doceniając ich aktywność podczas szkoły.

http://wwcs2019.org/

PROJEKTY WARSZTATOWE

Zaproponowano ponad 25 projektów naukowych, a ostatecznie do realizacji przez uczestników wybranych zostało 13, one również były bardzo interdyscyplinarne. Badaliśmy więc m.in.

- propagację opóźnień w sieci kolejowej;
- „społeczny metabolizm";
- zmiany położenia piłkarzy i ich wzorce, w trakcie trwania meczu piłki nożnej;
- stabilność ekologicznych sieci pokarmowych;
- metody predykcji samobójstwa na podstawie analizy tekstu autora;
- dynamikę cen nieruchomości w Londynie;
- matematyczny model uzależnienia od opiatów;
- propagację memów na portalach społecznościowych.

↓ Uczestnicy Zimowych Warsztatów Układów Złożonych

Celem Uczelnianej Oferty Dydaktycznej Centrum Studiów Zaawansowanych PW (UOD CSZ PW) jest poszerzanie wiedzy w wybranych kierunkach, a także pomoc i inspiracja w planowanej działalności naukowej. Program oferty adresowany jest do całego środowiska akademickiego Politechniki Warszawskiej, oraz chętnych spoza Uczelni. Na propozycję UOD CSZ PW składają się m.in. cykle interdyscyplinarnych wykładów podstawowych i specjalnych.

Merytoryczną opiekę nad UOD CSZ PW sprawuje Rada Programowa Centrum, którą tworzą naukowcy z Politechniki Warszawskiej, Uniwersytetu Warszawskiego, a także Polskiej Akademii Nauk.

Uczelniana Oferta Dydaktyczna Centrum Studiów Zaawansowanych

2019/2020

wykłady podstawowe (30 h)

Z1: **Modele matematyczne procesów i przemian** - prof. Stanisław Janeczko (PW)

Z2: **Równania różniczkowe cząstkowe dla inżynierów** - dr hab. Tomasz Cieślak (PAN)

Z3: **Matematyczna struktura czasu i przestrzeni I. Wstęp do szczególnej teorii względności** - prof. dr hab. Jerzy Kijowski (PAN)

Z4: **Wprowadzenie do uczenia maszyn** - prof. dr hab. inż. Władysław Homenda (PW)

Z6: **Niezwykle szczególna teoria względności** - dr Andrzej Dragan (UW)

L1: **Matematyczna struktura czasu i przestrzeni II. Wstęp do ogólnej teorii względności** - prof. dr hab. Jerzy Kijowski (PAN)

L3: **Teoria drgań** - prof. Piotr Przybyłowicz (PW)

L4: **Geny, GMO i Genetyka** - prof. Ewa Bartnik (UW)

wykłady specjalne (15 lub 30 h)

SZ1: **Analiza regresji w praktyce** (30h) - dr hab. inż. Anna Dembińska (PW)

SZ2: **Współczesne zarządzanie dla inżynierów i naukowców** (30h) - dr hab. inż. Janusz Zawiła-Niedźwiecki (PW)

SZ4: **Metody prowadzenia badań i statystycznej analizy wyników** (30h) - prof. Marek Dobosz (PW)

SZ5: **Etyka a teoria ewolucji** - dr hab. Adrian Kuźniar (UW)

SZ6: **Psychologia biznesu dla inżynierów** (30h) - wykład bez punktow ECTS dr Leszek Mellibruda (Active Business Mind Psychologia biznesu)

SZ7: **Iluzja Wiedzy i Granice Poznania** - prof. Stanisław Janeczko (PW)

SZ8: **Metody teorii osobliwości symplektycznych** - prof. Stanisław Janeczko (PW)

SL3: **Wybrane zagadnienia termodynamiki technicznej** (30h) - prof. Tomasz Wiśniewski (PW)

SL4: **Jak wydobyć potencjał twórczy grupy? Techniki pracy twórczej w grupie** dr Bartłomiej Skowron (PW)

SL5: **Psychologia zarządzania dla inżynierów** (30h) - wykład bez punktow ECTS dr Leszek Mellibruda (Active Business Mind Psychologia biznesu)

Uaktualniona lista przedmiotów znajduje się na stronie internetowej Centrum
wykłady podstawowe: http://www.konwersatorium.pw.edu.pl/oferta/w_podstawowe.html
wykłady specjalne: http://www.konwersatorium.pw.edu.pl/oferta/w_specjalne.html

Biuletyn Centrum Studiów Zaawansowanych „Profundere Scientiam"
Pl. Politechniki 1, p.152-154, 00-661 Warszawa; e-mail: csz@pw.edu.pl, www.csz.pw.edu.pl

Zespół redakcyjny: Małgorzata Zielińska – redaktor naczelna, Jowita Krakowiecka, Ilona Sadowska
Opieka merytoryczna: prof. Stanisław Janeczko
Projekt graficzny: Emilia Bojańczyk / Podpunkt | Opracowanie i skład: Małgorzata Zielińska / CSZ