

Piotr Giza

## **Maszyny a metoda naukowa**

### **1. WSTĘP**

Celem niniejszej pracy jest przedstawienie wpływu pewnych współczesnych osiągnięć w dziedzinie sztucznej inteligencji na niektóre aspekty niegdyś gorąco dyskutowanego, a obecnie już przebrzmiałego sporu między indukcjonizmem Franciszka Bacona a falsyfikacjonizmem Karla Poppera. Brytyjski filozof nauki i historyk matematyki Donald Gillies (1996) twierdzi na podstawie analizy historycznych przypadków odkryć, że metoda mechanicznej indukcji, którą opisuje i broni Bacon w *Nowym Organonie*, nie była jak dotąd prawie wcale stosowana w nauce. Argumentuje jednak, że sytuacja ta zmieniła się dzięki rozwojowi gałęzi sztucznej inteligencji zwanej *machine learning*, która prawdopodobnie przyniesie głębokie zmiany w sposobie uprawiania nauki. Metoda, którą można by określić jako indukcję bacońską, istnieje i zaczyna być regularnie stosowana.

Oczywiście spór o metodę naukową jest współcześnie niezwykle rozbudowany i bynajmniej nie toczy się między zwolennikami Bacona i Poppera. Oba stanowiska poddano krytyce oraz istotnym modyfikacjom. I tak, indukcjonizm Bacona nie uwzględniał wielu założeń metodologicznych i filozoficznych koniecznych do wyboru i eliminacji hipotez ani też faktu, że hipotezy niosą w sobie już pewną wiedzę tła. Z kolei falsyfikacjonizm miał poważne problemy z określeniem, co i w jakich sytuacjach należy odrzucić — hipotezę czy wiedzę towarzyszącą — a to groziło oskarżeniem o konwencjonalizm. Znane próby uratowania sytuacji podjął Lakatos, który niestety przed śmiercią nie zdążył odpowiedzieć na postawione przez Feysera zarzuty o anarchizm epistemologiczny.

Zarówno indukcjonizm, jak i falsyfikacjonizm napotkały poważne trudności w poszukiwaniu logiki nauki, nie uwzględniały one jednak faktu, że logika nie koń-

czy się na klasycznym rachunku predykatów. W tym kontekście bardzo obiecujące jest podejście polskich badaczy oparte na logice erotetycznej, między innymi dlatego, że pozwala reprezentować sytuację problemową mogącą prowadzić do rewizji wiedzy zastanej.<sup>1</sup>

Jeśli zatem w dalszych częściach pracy porównuję, biorąc za punkt wyjścia rozważania Gilliesa, pewne elementy stanowisk Bacona i Poppera, to czynię to z pełną świadomością, że oba te stanowiska od dawna nie są już reprezentatywne ani dla filozoficznego sporu o metodę naukową czy logikę nauki, ani nawet dla nurtów indukcjonizmu i falsyfikacjonizmu. Chodzi mi natomiast o wydobycie pewnych charakterystycznych ich aspektów czy rysów, a mianowicie skrajnie różniących się odpowiedzi, których udzielają na pytanie: skąd biorą się hipotezy czy twierdzenia naukowe?

Rozpaczynam od krótkiego przedstawienia, odpowiednio w sekcjach 2.1 i 2.2, indukcjonizmu Franciszka Bacona oraz falsyfikacjonizmu Karla Poppera. Popper w sposób radykalny odrzuca pogląd Bacona, jakoby istniała metoda indukcji pozwalająca wywnioskować prawa naukowe z dużej liczby opisów przypadków jednostkowych i proponuje całkowicie odmienny pogląd na kwestię, jak powinny przebiegać badania naukowe. Krytykując indukcję jako metodę nauki, stwierdza bardzo mocno, że w nauce (a także w życiu codziennym) wnioskowanie oparte na wielu obserwacjach jest mitem. Gillies argumentuje jednak, że metoda polegająca na mechanicznym generowaniu i eliminowaniu hipotez, którą można by określić jako indukcję baconowską, zaczyna być regularnie stosowana w komputerowych systemach *machine learning*. Z drugiej jednak strony, w funkcjonowaniu systemów *machine learning* występują także istotne elementy metodologii nauk Karla Poppera z jego naciskiem na procesy falsyfikacji hipotez. Pozwala to dostrzec fakt, że mimo oczywistych rozbieżności w innych kwestiach poglądy obu filozofów na rolę falsyfikacji w zdobywaniu wiedzy są w dużej części zgodne.

W sekcji 3.1 postaram się wyciągnąć z osiągnięć w dziedzinie *machine learning* pewne wnioski filozoficzne dotyczące sporu indukcjonizmu z falsyfikacjonizmem. Poddam też krytycznej analizie wnioski samego Gilliesa, które — jak mi się wydaje — są zbyt daleko idące, a jego ocena osiągnięć systemów *machine learning* nieco zbyt entuzjastyczna. Okazuje się jednak, że Gillies dokonuje niepełnej analizy stanu badań w dziedzinie odkryć maszynowych. Pomija bowiem niektóre ważne metodologicznie, choć mniej znane, osiągnięcia grupy badaczy skupionych wokół Herberta Simona oraz bardzo obiecujący program badawczy w dziedzinie maszynowych symulacji procesów poznawczych, zwany od nazwisk jego autorów HHNT. Osiągnięcia tego ostatniego programu zdają się wskazywać, że indukcja stanowi podstawową metodą zdobywania nowej wiedzy, także tej o charakterze teoretycznym. Dwa wspomniane programy badawcze oraz ich możliwe implikacje dla sporu indukcjonizmu z falsyfikacjonizmem omawiam odpowiednio w sekcjach 3.2 i 3.3. Dokładna

---

<sup>1</sup> Patrz np. Grobler 2006, s. 91nn. Praca ta zawiera też, w rozdziale I, znakomite streszczenie filozoficznego sporu o metodę naukową.

analiza osiągnięć programu badawczego HHNT znacznie wykraczałaby jednak poza ramy niniejszej pracy, tym bardziej że program ten nadal się rozwija i na filozoficzną czy metodologiczną ocenę wielu wyników jest, moim zdaniem, za wcześnie. Niemniej postaram się uwzględnić jego osiągnięcia w odniesieniu do podstawowej, postawionej przez Gilliesa kwestii: tego, czy (i jakiego rodzaju) indukcja stała się elementem metody naukowej.

## 2. INDUKCJONIZM FRANCISZKA BACONA A FALSZYFIKACJONIZM KARLA POPPERA

### 2.1. Franciszek Bacon

Indukcjonistyczna koncepcja metody naukowej została po raz pierwszy sformułowana przez Bacona, przede wszystkim w pracy *Novum Organum* z 1620 roku. W zamierzeniu autora miała to być systematyczna, niezawodna metoda zdobywania nowej, wartościowej i *prawdziwej* wiedzy naukowej. Bacon miał nadzieję, że gdy zastosuje się jego mechaniczne metody zdobywania wiedzy, nauka stanie się działalnością rutynową, niewymagającą specjalnej inteligencji czy pomysłowości. Formuluje swe idee za pomocą szeregu kwiecistych porównań i analogii charakterystycznych dla Anglii czasów Szekspira. Powiada na przykład, że o ile odręczne narysowanie linii prostej czy koła wymaga pewnej i wprawnej ręki artysty, o tyle uczynienie tego za pomocą linijki czy cyrkla takich umiejętności nie wymaga (1955, s. 80). Właśnie znalezienie tego rodzaju narzędzi było jego celem.

Podstawowa idea indukcjonizmu była następująca: naukowiec powinien rozpocząć od przeprowadzenia dużej liczby uważnych, szczegółowych obserwacji, by następnie z całej masy zebranych danych wyekstrahować prawa naukowe. Bacon (1955, s. 152) czyni nawet porównanie do ekstrakcji napoju winnego z wielu zebranych owoców winogron. Wierzył, że metoda ta okaże się o wiele bardziej owocna dla rozwoju nauki i technologii niż metoda sylogizmów Arystotelesa.

W *Nowym Organonie* Bacon ilustruje swą metodę podając szczegółową analizę mającą prowadzić do zbadania istoty czy, jak to określa, „formy” ciepła. Wyróżnia on w swej metodzie trzy kroki:

Pierwszy polega na „przygotowaniu *Historii Naturalnej i Eksperymentalnej*” (s. 174), co sprowadza się do zebrania obserwacji i przeprowadzenia eksperymentów związanych z badaną dziedziną.

Drugi krok jest konieczny dla uporządkowania i usystematyzowania zebranych danych i polega na utworzeniu tablic i zestawień faktów w taki sposób, aby rozum był w stanie się z nimi obchodzić (s. 174).

Dopiero na tym etapie, powiada Bacon, można zastosować krok trzeci: „[...] należy zastosować indukcję, właściwą i prawdziwą indukcję, która właśnie stanowi istotny klucz tłumaczenia” (s. 175).

Bacon rozpoczyna od zbudowania tablicy zwanej *tablicą istnienia i obecności* zawierającej 27 przypadków sytuacji, w których mamy do czynienia z ciepłem. Są to na przykład (s. 175 nn.):

- promienie Słońca, zwłaszcza latem i w południe,
- wszelkie postacie ognia,
- podgrzane lub gotujące się płyny,
- wszystkie ciała, niezależnie czy w postaci stałej, czy płynnej trzymane w pobliżu ognia,
- wszystkie ciała gwałtownie potarte,
- żelazo podczas rozpuszczania w kwasach,
- zwierzęta wytwarzające ciepło od wewnątrz.

Następnie opierając się o pierwszą tabelę, Bacon tworzy drugą, zawierającą przypadki, kiedy nie mamy do czynienia z ciepłem. Postępuje w następujący sposób: dla każdego przypadku z pierwszej tabeli stara się znaleźć przypadek jak najbardziej do niego podobny, w którym jednakże ciepło nie występuje. Tablicę tę nazywa *tablicą odchylenia, czyli nieobecności w przypadkach pokrewnych* (s. 178). I tak przykładowo dla pierwszego wyżej przytoczonego przypadku z poprzedniej tabeli stwierdza, że:

- promienie Księżyca, gwiazd ani komet nie dają odczucia ciepła.

W innych jednakże przypadkach wymienionych w powyższej tablicy Bacon nie dostrzega odpowiadających im przypadków braku występowania ciepła. Przykładowo, w odniesieniu do przypadku pocierania ciał z tablicy pierwszej stwierdza:

Sądźmy, że z tym wypadkiem pozytywnym nie łączy się negatywny. Nie znamy bowiem żadnego ciała dotykającego, które by na skutek tarcia widocznie nie stawało się ciepłe (s. 186).

Wreszcie trzecia tabela, zwana *tablicą stopni, czyli tablicą porównawczą odnoszącą się do ciepła* (s. 191), zawiera 41 przypadków pozwalających porównać intensywność ciepła w różnych sytuacjach. Przykładowo:

Ruch zwiększa ciepło, jak to zauważyć można przy miechu kowalskim i dmuchaniu: twardsze spośród metali nie rozpuszczają się ani nie topią pod działaniem ognia martwego albo spokojnego, jeśli nie podnieść go przez dmuchanie (s. 198).

Trzy opisane tabele łącznie określa Bacon jako *przedstawienie faktów rozumowi*.

Do tak uporządkowanych i usystematyzowanych danych stosuje Bacon swą metodę indukcji, by znaleźć *formę* ciepła. Należy, powiada, „[...] znaleźć taką własność, która by jednocześnie z daną własnością zawsze była obecna i nieobecna, wzrastała z nią razem lub malała, która by stanowiła szczególny przypadek własności bardziej ogólnej” (s. 203–204). W tym miejscu pada znamienne stwierdzenie:

Jeśli jednak umysł już od początku będzie usiłował to robić afirmatywnie (co zawsze zwykł czynić pozostawiony samemu sobie), wystąpią urojenia, przypuszczenia, pojęcia źle określone

i zasady wymagające codziennie poprawek [...] Odziałem samego Boga, może także aniołów i wyższych inteligencji jest znać formy bezpośrednio przez afirmację już od początku rozważania. Z pewnością jednak przekracza to możliwości człowieka, któremu dane jest postępować zrazu przez negację, a dopiero po zastosowaniu wszelkiego rodzaju wyłączeń zakończyć wreszcie afirmacją (s. 204).

Owe drogi afirmatywne i negatywne rozwiązywania problemów, które Bacon przeciwstawia sobie, to w bardziej współczesnym ujęciu metoda syntetyczna i analityczna.

Rozpoczyna więc od procesu, który określa jako *wykluczenie pewnych własności z formy ciepła*. Weźmy pewną dowolną własność ciepła A. Jeśli uda nam się znaleźć w pierwszej tabeli przypadek, gdzie występuje ciepło, lecz brak własności A, oznacza to, że występowanie ciepła nie pociąga za sobą A. Z drugiej strony, jeśli uda nam się znaleźć w drugiej tabeli przykład sytuacji, gdy występuje A, lecz nie mamy do czynienia z ciepłem — oznacza to, że występowanie A nie implikuje występowania ciepła.

Pierwszym więc zadaniem prawdziwej indukcji (jeśli chodzi o wykrycie form) jest odrzucenie, czyli wyłączenie poszczególnych własności, które nie występują w jakimś przypadku, gdzie występuje jedna własność, albo znajdują się w jakimś przypadku, w którym danej własności nie ma, albo też wrastają w jakimś przypadku, kiedy dana własność maleje, albo wreszcie maleją, kiedy dana własność wzrasta. Wtedy zaś po dokonaniu — jak należy — odrzuceniu i wyłączeniu [...] na drugim miejscu (jakby na dnie) pozostanie forma pozytywna, trwała, prawdziwa i dobrze określona (s. 205).

Stosując opisaną procedurę Bacon wyklucza i odrzuca świecenia i jasność, gdyż promienie Księżyca są jasne i świecące, lecz nie są ciepłe, podczas gdy żelazo może być bardzo gorące jednocześnie nie będąc jasnym i nie świecąc.

Bacon miał prawdopodobnie nadzieję, aczkolwiek w tej kwestii istnieją różnice interpretacji jego tekstu (Gillies, 1996, s. 60), że opisany proces czysto mechanicznego wykluczania i odrzucania własności na podstawie przypadków zgromadzonych w tabelach pozwoli uzyskać jednoznaczny rezultat — prawdziwą formę ciepła. Tak się jednak nie stało, w związku z czym Bacon stosuje drogę afirmacji — subtelne rozważania teoretyczne mające wprowadzić element teorii” do jego mechanicznej indukcji. Czyni to z wahaniem, jak twierdzi np. Gillies (s. 60), niemniej musi to uczynić, aby otrzymać jednoznaczną, sławną dziś konkluzję:

[...] forma, czyli prawdziwa definicja ciepła [...] ujęta w kilku słowach [da się wyrazić następująco]: ciepło jest ruchem, ekspansywnym, skrępowanym i napierającym poprzez mniejsze cząstki ciał (s. 218).

Bacon, lecz jest to moje bardzo osobiste przypuszczenie, nieuzasadnione w sposób należyty właściwymi badaniami jego tekstów, stosuje metodę, którą dziś nazwalibyśmy analizą statystyczną. Na podstawie danych zebranych w tablicach, a także innych obserwacji i analogii, dopatruje się bowiem korelacji między ciepłem a ruchem, wykluczając przy tym ruch ciała jako całości, które występują w większości przypadków. Bardzo podobnie, jak postaram się pokazać w sekcji 3.1, postępują współczesne systemy *machine learning*.

## 2.2. Karl Popper

Popper w sposób radykalny odrzuca pogląd Bacona, jakoby istniała metoda indukcji pozwalająca wywnioskować prawa naukowe z dużej liczby opisów przypadków jednostkowych i proponuje całkowicie odmienny pogląd na kwestię, jak powinny przebiegać badania naukowe. Krytykując indukcję jako metodę nauki, pisze znane zdanie:

Indukcja, tzn. wnioskowanie oparte na wielu obserwacjach, jest mitem. Nie stanowi ona faktu psychologicznego, nie występuje też w zwykłym, codziennym życiu ani w procedurach stosowanych przez naukę (1963, s. 53),

uzasadniając to dalej głównie faktem, iż nie możemy obserwować bez pewnego tła czy podbudowy teorii.

Popper proponuje w zamian teorię nauki opartą na domysłach i refutacjach — falsyfikacjonizm. Głosi on, że nauka rozpoczyna się nie od obserwacji, jak twierdzi indukcyjnizm, lecz od domysłów (*conjectures*). Następnie naukowiec stara się hipotezę obalić, sfalsyfikować, przez krytyczną analizę, obserwacje i eksperymenty. Hipoteza, która przetrwała pewną liczbę wymagających testów, zostaje tymczasowo zaakceptowana. Jeśli jednak hipoteza zostanie sfalsyfikowana, naukowcy muszą ją odrzucić, zastępując zmodyfikowaną odpowiednio hipotezą lub całkowicie nową. Nauka rozwija się przez nigdy niekończący się ciąg tworzenia i odrzucania hipotez — twierdzi Popper.

Co więcej, Popper i Bacon różnią się co do tego, w jaki sposób powstają nowe hipotezy naukowe. Bacon miał nadzieję, że teorie naukowe mogą zostać utworzone czy sformułowane na podstawie obserwacji za pomocą pewnego rodzaju procesu o charakterze mechanicznym. Popper przeciwnie był przekonany, że teorie naukowe są produktem twórczego myślenia wybitnych naukowców, takich jak Einstein, opierających się na jakiejś tajemniczej intuicji, która być może mogłaby stać się przedmiotem badania psychologów, lecz z całą pewnością nie da się sprowadzić do procedur logicznych:

[...] praca naukowca polega na wysuwaniu i sprawdzaniu teorii. Wydaje mi się, że stadium początkowe, akt powzięcia pomysłu czy wymyślenia teorii, ani nie wymaga logicznej analizy, ani się takiej analizie nie poddaje. Pytanie, jak to się dzieje, że ktoś wpada na nowy pomysł — czy to będzie temat muzyczny, intryga dramatyczna czy teoria naukowa — może być niezmiernie interesujące dla psychologii empirycznej, jest jednak bez znaczenia dla logicznej analizy wiedzy naukowej. Analiza logiczna nie zajmuje się bowiem *pytaniami o fakty* [...] ale wyłącznie pytaniami dotyczącymi *prawomocności* lub *ważności* [...].

Wprowadzę przeto wyraźne rozróżnienie pomiędzy procesem rodzenia się nowego pomysłu a metodami i wynikami jego logicznej analizy. Jeśli chodzi o zadanie logiki wiedzy — w przeciwieństwie do psychologii wiedzy — oprę się na założeniu, że polega ona jedynie na badaniu metod stosowanych w trakcie systematycznego sprawdzania, jakiemu trzeba poddać każdą nową koncepcję, jeśli mamy ją poważnie wziąć pod uwagę.

Jednakże wedle mego stanowiska w tej sprawie [...] nie istnieje nic takiego, jak logiczna metoda wpadania na nowe pomysły lub logiczna rekonstrukcja owego procesu. Stanowisko swe ująć mogę, mówiąc, iż każde odkrycie kryje „element irracjonalny” albo „intuicję twórczą” w sensie Bergsona (Popper, 1977, s. 32–33).

Dla Poppera więc tajemnicza intuicja i twórcze myślenie są w takim samym stopniu potrzebne wybitnemu naukowcowi, jak i wybitnemu artyście. Tylko w ten bowiem sposób naukowcy mogą tworzyć nowe hipotezy, konieczne, aby nauka mogła się posuwać naprzód. Popper przeciwstawia się więc często głooszonym poglądom (być może pod wpływem Bacona), że w odróżnieniu od artystów naukowcy postępują w sposób bardziej mechaniczny, racjonalny i rutynowy.

### **3. SYSTEMY ODKRYĆ MASZYNOWYCH A SPÓR INDUKCJONIZM—FALSYFIKACJONIZM?**

Wspomniany już we wstępie brytyjski filozof nauki i matematyk Donald Gillies (1996, s. 17) twierdzi, na podstawie analizy historycznych przypadków odkryć, że metoda mechanicznej indukcji, którą opisuje i której broni Bacon w *Nowym Organonie*, nie była jak dotąd stosowana w nauce w czystej postaci. Nieliczne przypadki jej rzekomego stosowania, na które zwraca uwagę Gillies, ilustrują raczej wnioskowanie na podstawie reguł heurystycznych niż czystą indukcję z wcześniej zgromadzonych danych. Metoda ta została podjęta i rozwijana dopiero po trzystu pięćdziesięciu latach przez badaczy sztucznej inteligencji.

Rozważania Bacona miały niewątpliwie charakter przyszłościowy. Jego entuzjazm dla rozwoju techniki, nowych instrumentów otwierających i poszerzających horyzonty badań naukowych można, jak twierdzi Gillies, rozszerzyć także na rozwój instrumentów wspomagających myślenie naukowe. Bacon nie dysponował jednak żadnymi tego typu fizycznymi instrumentami z wyjątkiem ołówka i papieru. Z tego też powodu był zmuszony ograniczyć się jedynie do zaleceń dotyczących stosowania bardziej systematycznych procedur polegających na przykład na układaniu tablic różnic i podobieństw. Jednak procedury te, bez wsparcia w postaci mogącego je zastosować fizycznego urządzenia, miały niewielkie szanse powodzenia i nie wykazywały się wyższością w stosunku do mniej systematycznych, intuicyjnych rozumowań właściwych istotom ludzkim.

„Dopiero dzięki komputerom mechaniczna metoda rozumowania Bacona mogła zostać po raz pierwszy zrealizowana” — pisze Gillies (1996, s. 68). Rozwój instrumentów wspomagających ludzkie narządy zmysłów (takich jak teleskop czy mikroskop) zmienił sposób uprawiania nauki i przyczynił się do ogromnego wzrostu wiedzy. Czyż więc nie mamy prawa przypuszczać, stwierdza dalej autor, że rozwój instrumentów wspomagających ludzki umysł może mieć podobny wpływ na naukę? Powinniśmy bowiem dostrzec istotną okoliczność: metoda naukowa nie ma ustalonego raz na zawsze charakteru. Przeciwnie, rozwija się wraz z rozwojem samej na-

uki. Rozwój komputerów i dziedziny sztucznej inteligencji, zwanej *machine learning*, zmienił, zdaniem Gilliesa, sposób uprawiania nauki i sprawił, że indukcja baconowska stała się standardową procedurą stosowaną w nauce.

Tradycyjne próby rozstrzygnięcia sporu pomiędzy indukcjonizmem a falsyfikacjonizmem sprowadzają się do analizy konkretnych historycznych przypadków odkryć naukowych po to, by odpowiedzieć na pytanie, który z konkurencyjnych modeli metody naukowej lepiej je wyjaśnia. Sądzę, że wraz z powstaniem komputerowych systemów odkryć naukowych filozofowie nauki weszli w posiadanie nowego narzędzia do empirycznego testowania swych koncepcji nauki: miarą, mianowicie sukcesu czy porażki jakiejś koncepcji rozwoju nauki może być to, w jakim stopniu systemy odkryć naukowych, które stosują tę metodę, „radzą sobie” z dokonywaniem odkryć naukowych.

W sekcji 3.1 przedstawię nieco krytyczną analizę osiągnięć i niedostatków systemów *machine learning* w dziedzinie odkryć naukowych. W kolejnej sekcji 3.2 zaprezentuję ważne, lecz pominięte przez Gilliesa osiągnięcia chyba najbardziej znanego programu badawczego w dziedzinie odkryć maszynowych, sformułowanego przez tzw. grupę Herberta Simona. Wreszcie w sekcji 3.3 postaram się wyciągnąć pewne wnioski filozoficzne dotyczące metody naukowej z osiągnięć zapewne najbardziej obiecującego programu badawczego zwanego od nazwisk swych twórców HHNT.

### 3.1 Systemy *machine learning* w tradycji Alana Turinga

Systemy *machine learning* w tradycji Turinga historycznie wywodzą się od jego prac nad koncepcją maszyny nazwanej potem jego nazwiskiem oraz jego prac w dziedzinie kryptologii prowadzonych podczas II wojny światowej wraz z Donaldem Michie w ośrodku badawczym w Bletchley w Anglii. To właśnie Michie przyczynił się po wojnie do rozpowszechnienia idei Turinga w Europie i poza nią.

Kluczowymi cechami wyróżniającymi tę tradycję badawczą jest użycie aparatury logicznej i języka programowania Prolog do rozwiązywania konkretnych problemów oraz nastawienie na praktyczne zastosowania systemów w wielu dziedzinach w przemyśle, medycynie czy nauce. Podkreślić przy tym należy, że autorzy systemów nigdy nie byli bezpośrednio zaangażowani w spory filozoficzne czy metodologiczne dotyczące metody naukowej, a konsekwencje osiągnięć tych systemów dla sporu dotyczącego metody naukowej dostrzegł dopiero Donald Gillies.

Systemy *machine learning* zostały stworzone po to, aby pomóc pokonać trudności, które pojawiły się w tworzeniu i wdrażaniu systemów eksperckich, czyli systemów zaprojektowanych po to, aby wykonywały zadania, które normalnie wykonują eksperci w przemyśle, medycynie czy nauce. Pierwszym udanym systemem eksperckim był DENDRAL (Buchanan i Feigenbaum, 1978), stworzony na uniwersytecie Stanford w 1965 roku w ramach tzw. projektu programowania heurystycznego przez E. Feigenbauma i B. Buchanana we współpracy z grupą badaczy zajmujących się



analizą spektrogramów skomplikowanych związków organicznych. Już w roku 1975 kolejny system stworzony w ramach tego projektu MYCIN wypadł w testach lepiej niż 9 doświadczonych lekarzy-diagnostów.

W krótkim czasie stało się jasne, że systemy eksperckie odnoszą sukcesy i są przydatne w wielu dziedzinach, jednak na drodze ich dalszego rozwoju pojawiły się trudności z pozyskiwaniem wiedzy od ekspertów i formułowaniem na jej podstawie tzw. praw produkcji w bazie wiedzy systemu. Początkowo odbywało się to metodą długich, żmudnych wywiadów z ekspertami, wkrótce okazało się jednak, że w wielu przypadkach dokładne wyartykułowanie i zwerbalizowanie wiedzy jest dla ekspertów trudne, a jeszcze trudniejsze lub wręcz niemożliwe jest jej sformułowanie w sposób na tyle ścisły i logiczny, aby mogła posłużyć do budowy bazy praw systemu eksperckiego. Tę istotną przeszkodę dla szybko rozwijającej się dziedziny badań i zastosowań praktycznych nazwano „wąskim gardłem Feigenbauma”.

Systemy *machine learning* stworzono właśnie w celu pokonania wspomnianych trudności: zamiast bowiem przeprowadzać żmudne i czasochłonne wywiady z ekspertami w celu sformułowania praw dla bazy wiedzy tworzonego systemu eksperckiego można zbudować system, który sformułuje je drogą indukcji na podstawie indywidualnych przypadków jednoznacznie zakwalifikowanych przez ekspertów. Nie znaczy to, że celem systemów *machine learning* jest całkowite wyeliminowanie ekspertów — przeciwnie, eksperci dostarczają tym systemom istotnej wiedzy towarzyszącej, formułują ważne przykłady, od których systemy rozpoczynają konstruowanie praw, wreszcie oceniają i korygują otrzymane przez systemy prawa. Pierwszym tego typu systemem był Meta-DENDRAL, którego wyniki w tworzeniu praw produkcji dla systemu DENDRAL były na tyle spektakularne, że opublikowano je w specjalistycznym czasopiśmie chemicznym (Gillies, 1996, s. 29; Buchanan and Feigenbaum, 1978). Gillies, w swoich filozoficznych analizach koncentruje się na nowszych systemach powstałych w tradycji Turinga, które, jak utrzymuje, mają istotne znaczenie dla sporu indukcyjności z falsyfikacjonizmem. Omawia on dokładnie dwa z nich, są to GOLEM i ID3.

ID3 (Quinlan, 1979, 1986) został prawdopodobnie wybrany przez Gilliesa z powodu swych walorów dydaktycznych, ponieważ bardzo dobrze ilustruje procedurę indukcyjną pozwalającą sformułować tzw. prawa klasyfikacyjne i drzewa decyzyjne na podstawie zbioru własności obiektów określonych za pomocą wielu przypadków jednostkowych. Jego działanie jest na tyle proste, by dało się dostrzec, że podstawowa reguła wnioskowania, stosowana przez system iteracyjnie, bardzo przypomina procedurę mechanicznej indukcji w duchu Bacona. Chociaż system posługuje się stosunkowo prostą logiką predykatów jednoargumentowych pierwszego rzędu, to, jak się okazuje, wystarcza mu to do rozwiązania wielu praktycznych problemów. Na wejściu otrzymuje on tzw. zbiór treningowy obiektów opisywanych przez skończoną liczbę predykatów oraz informację o tym, które obiekty są przypadkami pozytywnymi, a które negatywnymi pewnej określonej własności. Zadanie polega na sformułowaniu na drodze rozumowania indukcyjnego tzw. praw klasyfikacji obiektów

na podstawie ich własności. Prostim, podanym przez Gilliesa przykładem takiego prawa, klasyfikującego ptaki jako łabędzie i nie-łabędzie jest:

x jest łabędziem wtedy i tylko wtedy gdy x posiada rozmiar: duży, x posiada długość szyi: długa, x jest stworzeniem wodnym i x jest biały lub gdy x posiada rozmiar: duży, x posiada długość szyi: długa, x jest stworzeniem wodnym i x jest czarny (Gillies, 1996, s. 31).

przy założeniu, że wszystkie ptaki w zbiorze treningowym opisywane są przez atrybuty: rozmiar, długość szyi, bycie stworzeniem wodnym, kolor, z których każdy ma skończoną liczbę wartości oraz że każdy obiekt w tym zbiorze został jednoznacznie zakwalifikowany jako łabędź lub nie-łabędź.

System GOLEM (Muggleton and Feng, 1992) nie jest już tak prosty i intuicyjny jak ID3, stosuje bowiem wewnętrzną reprezentację danych opartą na rachunku predykatów pierwszego rzędu o większej niż jeden liczbie argumentów (relacjach), a jego zasada funkcjonowania, zwana uczeniem się relacyjnym, nadaje się do rozwiązywania problemów wykraczających poza możliwe zastosowania systemów odwołujących się jedynie do atrybutów obiektów. GOLEM używa tzw. logiki klauzul, charakterystycznej dla języka programowania logicznego PROLOG, który stanowi nie tylko narzędzie, ale także pewien sposób myślenia przy konstruowaniu systemów *machine learning*. Podobnie jak ID3, GOLEM rozpoczyna od zbioru treningowego obiektów, by następnie drogą stopniowych uogólnień indukcyjnych w kolejnych krokach iteracyjnych dojść do ogólnych praw klasyfikacji opartych na relacjach między obiektami i ich własnościami.

Bez nadmiernego wdawania się w szczegóły techniczne i różnice między poszczególnymi systemami należy stwierdzić, że Gillies (s. 18–19) eksponuje trzy istotne aspekty wspólne dla systemów *machine learning*:

1. Algorytmy systemów stosują mechaniczną metodę generowania hipotez na podstawie danych opartą na *indukcyjnych regułach wnioskowania*, które nie wymagają ludzkiej ingerencji ani pomysłowości i działają w duchu indukcji baronowskiej.

2. Reguła wnioskowania różni się nieco od oryginalnego wnioskowania indukcyjnego Bacona: hipoteza H nie wynika jedynie z danych E, lecz z koniunkcji danych i pewnej wiedzy tła czy wiedzy towarzyszącej, K, której dostarczają eksperci lub która jest wbudowana w reguły heurystyczne systemów. Jest to istotny element współpracy człowiek-komputer.

3. Systemy stosują swe podstawowe reguły wnioskowania w sposób iteracyjny, przy czym na każdym etapie istotną rolę odgrywają mechanizmy testowania i falsyfikacji hipotez w duchu metodologii Poppera, z tą tylko różnicą, że hipotezy generowane są mechanicznie przez system, a nie w sposób intuicyjny przez ludzi.

Zarówno ID3, jak i GOLEM okazały się bardzo przydatne przy formułowaniu praw dla systemów eksperckich w wielu dziedzinach zastosowań, między innymi w przemyśle i medycynie. Gillies jednak koncentruje się na pewnym bezpośrednim zastosowaniu systemu GOLEM w badaniach naukowych, twierdząc nieco na wyrost, że GOLEM dokonał odkrycia nowego prawa nauki o charakterze przyczynowym

(Gillies, 1996, s. 50), co więcej, odkrycie to związane jest z istotną i wciąż rozwijającą się dziedziną — badaniem struktury białek.

Gillies rozpoczyna swą analizę od przedstawienia stanu badań na temat aminokwasowej struktury białek na początku lat 90.<sup>2</sup> Białka składają się z 20 różnych aminokwasów, których łańcuchy połączone są ze sobą na końcach tak, że cząsteczka białka składa się z sekwencji rodników aminokwasowych. Ta sekwencja, zwana strukturą pierwszorzędową białek, jest stosunkowo łatwa do ustalenia, nie wystarczy jednak, niestety, by móc przewidzieć biologiczne własności białek. Zależą one bowiem w decydującym stopniu od kształtu przestrzennego ich cząsteczek, czyli tzw. struktury drugorzędowej. Można ją ustalić na drodze żmudnych, długich i kosztownych badań przy użyciu metod dyfrakcji promieni X lub magnetycznego rezonansu jądrowego. W cząsteczkach białek występują głównie dwie struktury przestrzenne, tzw. helisy  $\alpha$  (odkryte przez słynnego Linusa Paulinga już w 1951 roku) i harmonijki  $\beta$ . Białka można więc zaklasyfikować do 3 grup: 1) typu  $\alpha$ , w którym cząsteczki składają się jedynie z helisy  $\alpha$ , 2) typu  $\beta$ , w którym cząsteczki składają się jedynie z harmonijek  $\beta$ , oraz 3) typu mieszanego, gdzie występują obie struktury.

Postęp w dziedzinie biochemii byłby znacznie szybszy i łatwiejszy, gdybyśmy nauczyli się przewidywać strukturę drugorzędową białek na podstawie ich sekwencji aminokwasów, a więc struktury pierwszorzędowej. Niestety problem ten, mimo wieloletnich wysiłków uczonych, udało się rozwiązać jedynie w niewielkiej części. Najprostszą wersją tego problemu polega na ograniczeniu się jedynie do białek typu  $\alpha$  i próbie znalezienia odpowiedzi na pytanie, czy przy określonej znanej strukturze pierwszorzędowej dany rodnik tworzy strukturę  $\alpha$ , czy też jej nie tworzy. Do tak zawężonego, lecz nadal bardzo trudnego problemu zastosowano właśnie system GOLEM.

System otrzymał najpierw na wejściu zbiór treningowy składający się z dwunastu cząsteczek białka o znanej homologicznej strukturze typu  $\alpha$ , zawierających ponad 1500 rodników aminokwasowych. Na podstawie tego zbioru oraz wiedzy tła GOLEM sformułował pewną niewielką liczbę praw pozwalających przewidzieć, które rodniki są częściami helisy  $\alpha$ . Dane na temat przypadków pozytywnych i negatywnych zakodowane były w postaci predykatów dwuargumentowych o ogólnej postaci:  $\alpha(\text{nazwa\_białka}, \text{pozycja})$ , orzekających, że rodnik na określonej pozycji w łańcuchu danej cząsteczki białka tworzy helisę  $\alpha$ .

Istotną kwestią jest sposób reprezentacji wiedzy tła, stanowiącej zarazem, jak podkreśla Gillies, reguły heurystyczne systemu wyznaczające przestrzeń przeszukiwań. Rodniki białkowe opisywane są w systemie za pomocą takich, a nie innych własności uznanych, w świetle wiedzy specjalistów w danej dziedzinie, za istotne dla ich struktury drugorzędowej. Są to predykaty jednoargumentowe, orzekające np. o rozmiarze cząsteczki, własnościach hydrofobowych, biegunowości czy budowie aromatycznej, ale także dwuargumentowe, orzekające o relacjach między własnościami różnych cząsteczek. Inne predykaty w jawny sposób pozwalają systemowi

<sup>2</sup> Autor opiera się na pracy Muggleton, King i Sternberg, 1992.

wytyczyć najbardziej obiecujący kierunek poszukiwań regularności na podstawie wiedzy już zgromadzonej przez specjalistów na temat związków określonych sekwencji rodników w łańcuchu białek i ich struktury przestrzennej. Istotne ograniczenie polegające na tym, że wiedza tła systemu może mieć jedynie postać zdań jednostkowych, nie stanowi w przypadku badania struktury białek istotnej przeszkody — ich cząsteczki składają się ze skończonej niewielkiej liczby rodników, dzięki czemu istotna wiedza o ich znanych sekwencjach jest możliwa do zakodowania bez odwoływania się do formuł ogólnych.

GOLEM, podobnie jak inne systemy *machine learning*, poszukuje praw, stosując procedurę iteracyjną: prawa otrzymane przy jego pierwszym przebiegu zostały następnie dołączone do wiedzy tła, by w następnym przebiegu wygenerować prawa bardziej ogólne. Otrzymane przez system prawa były następnie testowane na czterech różnych białkach o znanej niehomologicznej strukturze należących do typu  $\alpha$ . Testy wykazały zgodność tych praw z danymi na poziomie 81% (Gillies, 1996, s. 51).

Gillies (1996, s. 53) przytacza jedno z tych praw (otrzymane przy pierwszym przebiegu systemu, a więc bezpośrednio na podstawie danych), które po przetłumaczeniu z języka klauzul Prologu na język naturalny brzmi następująco:

W cząsteczce białka A na pozycji B występuje rodnik tworzący helisę  $\alpha$ , jeżeli spełnione są następujące warunki:

- rodnik na pozycji B-2 nie jest proliną
- rodnik na pozycji B-1 nie jest aromatyczny ani nie jest proliną
- rodnik na pozycji B jest duży, niearomatyczny i nie jest lizyną
- rodnik na pozycji B+1 ma własności hydrofobowe i nie jest lizyną
- rodnik na pozycji B+2 nie jest aromatyczny ani nie jest proliną
- rodnik na pozycji B+3 nie jest aromatyczny ani nie jest proliną, ani nie jest mały, ani nie ma budowy biegunowej
- rodnik na pozycji B+4 ma własności hydrofobowe i nie jest lizyną.

Z powyżej przytoczonego przykładu Gillies wyciąga dwa dosyć mocne wnioski:

Stwierdza, że przytoczone prawo (i inne jemu podobne) jest całkowicie nowym prawem, które nie było znane aż do czasu sformułowania go przez system GOLEM. Co więcej, prawo to stanowi przyczynek do ważkich współczesnych badań nad przyrodą. Stąd też, jak utrzymuje (1996, s. 53), mamy prawo twierdzić, że system GOLEM *odkrył nowe prawo przyrody*.

Stwierdza za autorami systemu, że prawa tego typu mają charakter przyczynowy: występowanie konkretnych sekwencji rodników o określonych własnościach *powoduje* taką, a nie inną strukturę przestrzenną.

Uważam, że bardzo kompetentna i wartościowa analiza dokonana przez Gillies pokazuje, że sposób uprawiania nauki istotnie zmienił się od czasu, gdy Karl Popper pisał swe krytyczne słowa na temat indukcji, które przytoczyłem na początku sekcji 2.2 — metoda naukowa ewoluuje, podobnie jak sama nauka. Wraz z pojawieniem się i rozwojem systemów *machine learning* indukcja baconowska w nieco zmodyfi-

kowanej postaci, uwzględniająca także niektóre idee Poppera, stała się częścią procedur naukowych. Analiza ta pokazuje także ważne aspekty interakcji między komputerem a człowiekiem w badaniach naukowych.

Nie znaczy to jednak, że zgadzam się ze wszystkimi twierdzeniami Gilliesa. Przede wszystkim, można podać w wątpliwość stwierdzenie, czy prawa odkryte przez system GOLEM zasługują na miano *praw nauki*, a tym bardziej *przyczynowych praw nauki*. Wydaje mi się bowiem, że prawo w rodzaju przytoczonego wyżej jest co najwyżej uogólnieniem empirycznym niskiego poziomu o charakterze statystycznym, prawdziwym jedynie w około 80% przypadków. Po drugie, wbrew twierdzeniom Gilliesa nie uważałbym go za prawo przyczynowe: nie *wyjaśnia* ono mechanizmów przyczynowych powodujących, że określone własności pobliskich rodników w łańcuchu *powodują* powstanie struktury typu  $\alpha$  w danym rodniku, a jedynie stwierdza pewną korelację statystyczną.

Tak więc, podsumowując rozważania tej sekcji, należy stwierdzić, że dzięki szybkim komputerom i systemom *machine learning* indukcja baconowska z elementami falsyfikacji hipotez Poppera stała się częścią metody naukowej. Ogranicza się ona jednak do formułowania praw empirycznych o niskim poziomie ogólności bez elementów „teorii” i wyjaśnienia teoretycznego.

### 3.2. Grupa Herberta Simona

Najbardziej znanym programem badawczym, którego celem jest racjonalna rekonstrukcja i komputerowa symulacja procesu odkrycia naukowego, jest program sformułowany przez grupę badaczy skupionych wokół Herberta Simona. Choć pierwsze publikacje Simona poświęcone odkryciom naukowym ukazały się już w latach 60. XX wieku, dyskusje wokół „programu Simona” odwołują się zwykle do wspólnej książki opublikowanej w 1987 roku pod tytułem *Scientific Discovery: Computational Explorations of the Creative Process*, której autorami byli Pat Langley, Herbert Simon, Gary Bradshaw i Jan Żytkow. W odróżnieniu od pracujących w tradycji Turinga badaczy, których głównie interesowały zastosowania praktyczne, podejście grupy Simona ma głęboką podbudowę filozoficzną i metodologiczną. Już we wstępnym rozdziale wspomnianej książki znajdujemy następujący fragment:

Centralną hipotezą w naszych badaniach jest twierdzenie, że mechanizmy odkrycia naukowego nie mają specyficznego, unikalnego dla tej działalności charakteru, lecz są szczególnym przypadkiem ogólnych mechanizmów rozwiązywania problemów. [...] W naszych badaniach posłużymy się metodą polegającą na zbudowaniu programu komputerowego (a dokładnie serii takich programów) zdolnego do dokonywania nietrywialnych odkryć naukowych, którego zasada działania oparta jest na naszej wiedzy o mechanizmach rozwiązywania problemów przez ludzi [...] (Langley *et. al.* 1987, s. 5).

Używając heurystyk wzorowanych na zachowaniu ludzi badanych w laboratoriach psychologicznych oraz relacji dotyczących wybitnych odkryć w naukach przy-

rodniczych, grupa dokonała racjonalnej rekonstrukcji pojęciowej, a następnie zbudowała systemy komputerowe symulujące procesy odkrycia naukowego w różnych obszarach działalności badawczej, takich jak znajdowanie i formułowanie problemów, przeprowadzanie eksperymentów, formułowanie praw empirycznych na podstawie danych czy odkrywanie „ukrytej” struktury materii. Grupa rozwija swe systemy w najstarszym języku programowania sztucznej inteligencji — LISP, nadal powszechnie używanym w Stanach Zjednoczonych i Kanadzie.

Najbardziej znanym i najszerzej dyskutowanym programem stworzonym przez grupę jest system BACON, w istocie seria systemów od BACON.1 do BACON.5, z których każdy wprowadza nowe możliwości w symulowaniu odkrywania praw empirycznych na podstawie danych. Programy z serii BACON, startując od tych samych, wedle opinii ich autorów, warunków początkowych, z którymi mieli do czynienia prawdziwi odkrywcy, dokonały ponownego odkrycia wielu znanych praw z historii fizyki, jak trzecie prawo Keplera opisujące ruch planet, prawo bilansu cieplnego Blacka (z uwzględnieniem możliwych przejść fazowych), pierwsze i drugie prawo Ohma związane z oporem elektrycznym, prawo załamania światła Snelliusa i wiele innych. Autorzy twierdzą ponadto, że najbardziej rozbudowane wersje systemu BACON postulują, czy wprowadzają „terminy teoretyczne”, jak ciepło właściwe, masa bezwładna, współczynnik załamania (Langley *et al.* 1987, s. 66–86).

Obok ogólnych metod rozwiązywania problemów, jak heurystyczne przeszukiwanie selektywne, tzw. metoda *means-ends-analysis* czy techniki oparte na systemie produkcji (*production system*) i metodzie dopasowywania do wzorca (*pattern-matching*),<sup>3</sup> BACON i inne systemy dokonujące odkryć empirycznych używają prostych heurystyk sterowanych przez dane wejściowe, specyficznych dla procesu odkrywania praw empirycznych (Langley i Żytkow, 1989):

— Reguła 1: Jeśli dla dostępnych wartości  $X$  zmienna  $Y$  ma w przybliżeniu stałą wartość, to wyciągnij wniosek, że  $Y$  ma zawsze stałą wartość.

— Reguła 2: Jeśli dostępne wartości  $Y$  i  $X$  są związane liniowo, to przyjmij, że zawsze obowiązuje między nimi zależność liniowa.

— Reguła 3: Jeśli wraz ze wzrostem  $Y$  wzrasta  $X$ , zdefiniuj stosunek  $X/Y$  i zbadaj, jak zachowują się jego wartości.

— Reguła 4: Jeśli wraz ze wzrostem  $Y$  maleje  $X$ , zdefiniuj iloczyn  $XY$  i zbadaj, jak zachowują się jego wartości.

Te proste reguły heurystyczne okazały się wystarczające, aby systemy z serii BACON były w stanie dokonać spektakularnych odkryć praw wspomnianych wyżej. Dalsza szczegółowa analiza programu Simona i jego osiągnięć znacznie jednak przekraczałyby ramy niniejszego artykułu, którego głównym celem jest odpowiedź

<sup>3</sup> Patrz na przykład Langley *et al.*, 1987, rozdz. 1; Simon, 1992, s. 5–7.

na pytanie, czy i w jakim stopniu osiągnięcia systemów odkryć mają wpływ na rozstrzygnięcie sporu indukcyjnizm-falsyfikacjonizm, a ogólniej na metodę naukową.<sup>4</sup>

Mimo znaczących sukcesów w dziedzinie pojęciowej analizy procesów odkrycia naukowego w terminach rozwiązywania problemów oraz wielu wdrożonym systemom odkryć w różnych dziedzinach nauki, program Simona spotkał się z ostrą krytyką.<sup>5</sup> Większość krytyków, w tym także Donald Gillies, zarzuca badaczom z grupy Simona, że ich systemy nie dokonują (nawet *ponownie*) tych samych odkryć w *tych samych warunkach początkowych*, z którymi mieli do czynienia oryginalni odkrywcy. Zazwyczaj problemy w nauce nie są w pełni zdefiniowane (*ill-defined*) w odróżnieniu od typowych, nawet trudnych problemów stawianych przed ludźmi badanymi w laboratoriach psychologicznych czy problemów występujących w grach — powiadają krytycy. To właśnie dlatego ich rozwiązanie wymaga często wielu lat pracy wybitnych umysłów, podczas gdy problemy, z jakimi mają do czynienia systemy grupy Simona, są już precyzyjnie zdefiniowane, a komputer rozwiązuje je w ciągu minut, a nawet sekund. Tak sformułowany problem przypomina raczej podręcznikowe zadanie dla studentów niż prawdziwe odkrycie naukowe. Dlatego też nie powinno nas dziwić, że systemy Simona nie dokonały żadnego *nowego* odkrycia, stwierdza na przykład Gillies (1996, s. 24).

Aby zilustrować tego typu zarzuty rozważmy przykład często przywoływanego przez krytyków Simona trzeciego prawa Keplera:  $\frac{R^3}{T^2} = const$ , gdzie R oznacza średnią odległość planety od Słońca, a T okres jej obiegu. Odkrycie tego prawa było łatwe dla systemu BACON, ponieważ problem został precyzyjnie sformułowany i przygotowany przez autorów systemu. System „wiedział”, które zmienne są istotne dla problemu, ponieważ to właśnie ich wartości, znane także Keplerowi, otrzymał na wejściu, „wiedział” także, jakiego rodzaju zależności między nimi ma szukać. W odróżnieniu od tego problem, z którym musiał się zmierzyć Kepler, był trudny: nie wiedział on, jakie zmienne są istotne dla opisu ruchu planet (masa, kolor, jasność planety itd.) ani jakiego rodzaju zależności między nimi szukać. Co więcej, wybór zmiennych określających odległość planety i okres jej obiegu wokół Słońca, a nie zmiennych opisujących jej ruch względem Ziemi wydawał się mało spójny z obowiązującym w czasach Keplera obrazem świata (Gillies, 1992, s. 30).

Podejście grupy skupionej wokół Herberta Simona bywa określane jako „psychologiczne” w odróżnieniu od „logicznego” podejścia badaczy pracujących w tradycji Alana Turinga. Zdaniem Gilliesa jest to poważna wada metodologii Simona. Badacze z tej grupy rozpoczynają od analizy rozumowania, jakie wedle dostępnych relacji przeprowadzili znani odkrywcy, by następnie symulować to odkrycie za po-

<sup>4</sup> Obszerną krytyczną analizę programu badawczego grupy Simona znaleźć można na przykład w pracy Giza, 2006, rozdz. 2.2.1.

<sup>5</sup> Patrz np. specjalny numer *International Studies in the Philosophy of Science* (1992, vol. 6, issue 1) poświęcony teorii odkryć maszynowych H. Simona.

mocą programów komputerowych. Badacze pracujący w tradycji Turinga budują natomiast reguły inferencji swoich systemów na drodze analizy logicznej, bez odwoływania się do sposobu, w jaki analogiczne problemy rozwiązują ludzie. Analiza ważnych odkryć z historii nauki może nas wiele nauczyć na temat roli heurystyk czy wiedzy towarzyszącej w formułowaniu i ocenie hipotez naukowych, nie prowadzi ona jednak do formułowania ścisłych praw indukcyjnych, co zdaniem Gilliesa (1996, s. 23) stanowi główną przyczynę faktu, że systemy Simona nie są w stanie dokonać żadnego nowego odkrycia.

Otóż zdecydowanie nie zgadzam się z takim uogólnianiem. Istotnie, można zarzucać systemom, takim jak BACON, formułującym prawa empiryczne na podstawie danych, że ich autorzy wzorują się na historycznych zapisach rozumowań prowadzących do odkrycia danego prawa, a te z punktu widzenia logiki wnioskowania posiadają luki, nie do końca uprawnione przeskoky czy skróty myślowe. Krytycy jednak, w tym także Gillies, pomijają w swoich analizach systemy grupy Simona, których zadaniem było odkrycie „ukrytej” struktury materii w fizyce i chemii, takie jak STAHL i DALTON (Langley *et al.*, 1987), REVOLVER (Rose, 1988) czy GELL-MANN (Fisher i Żytkow, 1990). Wszystkie te systemy postulują pewne struktury niedostępne bezpośredniej obserwacji, w sposób mechaniczny generując hipotezy na drodze indukcji, a następnie weryfikując je i eliminując w duchu metodologii Poppera, by ograniczyć przestrzeń przeszukiwań tak wcześnie, jak to tylko możliwe.

W pracy poświęconej implikacjom komputerowych systemów odkryć naukowych dla problemu realizmu (Giza, 2002) analizuję dosyć szczegółowo najbardziej zaawansowany ze wspomnianych wyżej systemów, GELL-MANN, zaprojektowany w 1990 roku na uniwersytecie Wichita. Zadaniem systemu było dokonanie analizy danych dotyczących własności cząstek elementarnych (a dokładniej — tzw. hadronów) znanych w roku 1964, a następnie sformułowanie hipotezy (czy hipotez) na temat „ukrytej” w nich prostszej struktury materii lub mówiąc językiem współczesnym, odkrycie kwarków. We wspomnianej pracy pokazuję też, że nawet najbardziej zaawansowane systemy odkrywające „ukrytą” strukturę materii rozumują na innym poziomie teorii i używają innych metod niż oryginalni odkrywcy w analogicznej sytuacji problemowej. Te dwa poziomy korespondują z rozróżnieniem wprowadzonym przez Nancy Cartwright w głośnej swego czasu książce *How the Laws of Physics Lie* (1983) między teoriami i prawami fenomenologicznymi.

Wedle Cartwright zadaniem ogólnych teorii czy praw wysokiego poziomu (jak kwantowa teoria pola) jest *wyjaśnianie* szerokich grup zjawisk, co odbywa się kosztem dokładności opisu, podczas gdy prawa fenomenologiczne (jak prawo rozpadu promieniotwórczego) mają właśnie za zadanie dokładnie *opisywać* wąskie klasy zjawisk. Co więcej, Cartwright argumentuje, że owe dwa poziomy czy rodzaje praw funkcjonują w nauce w sposób w dużej mierze niezależny od siebie. Sądzę, że rozróżnienie wprowadzone przez Cartwright ma istotne znaczenie dla właściwej oceny metody stosowanej przez systemy poszukujące ukrytej struktury materii.



Otóż zamiast szukać najprostszej reprezentacji w ramach formalizmu relatywistycznego pola kwantowego i teorii grup, jak uczynili to fizycy-teoretycy w roku 1964, system GELL-MANN w sposób mechaniczny generuje możliwe modele kwarkowe, postulując liczbę typów „ukrytych” obiektów, liczbę składników każdej cząstki elementarnej w analizowanej grupie, wreszcie liczby kwantowe kwarków, a następnie dokonuje przeszukiwania ogromnej przestrzeni tych modeli, szukając najprostszego modelu zgodnego z danymi „obserwowalnymi” na temat cząstek elementarnych dostępnymi w roku 1964 (Giza, 2002; 2006, sekcja 3.2.3). System stosuje ostrożne uogólnienia indukcyjne, stopniowo starając się dopasować znalezione modele adekwatne dla opisu określonych rodzin hadronów, dla kolejnych znanych rodzin tych cząstek. Stosując jedynie arsenał teoretyczny oparty na prawach fenomenologicznych i własnościach hadronów znanych w roku 1964 oraz kryteria oparte na unikalności, prostocie i zgodności modelu z danymi, system GELL-MANN znajduje dokładnie jeden model, przy czym jest to właśnie model zaakceptowany przez fizyków.

Wydaje się oczywiste, że żaden poważny współczesny naukowiec nie postępowalby tak, jak uczynił to system GELL-MANN. Łatwo jest bowiem postulować ukrytą strukturę, ale bardzo trudno taką, a nie inną jej koncepcję uzasadnić — właśnie ze względu na wielość możliwych modeli. Jest to znany we współczesnej filozofii nauki problem *niedookreślenie teorii przez dane*. Naukowiec-człowiek nie mógłby nawet marzyć o uzasadnieniu jakiegokolwiek zaproponowanego przez siebie modelu wyłącznie na poziomie praw fenomenologicznych (w rozumieniu Cartwright), a więc bez odwoływania się do rozważań na poziomie ogólnych, abstrakcyjnych teorii. Aby tego dokonać, musiałby bowiem nie tylko pokazać, że jego model jest spójny z danymi, lecz także, że jest najprostszym spośród takich modeli. Odwołanie się do kryterium prostoty przez pokazanie, że nie istnieje żaden równie prosty lub prostszy model także spójny z danymi, wymaga jednak przeanalizowania ogromnej liczby modeli (w dyskutowanym przypadku liczba ta sięga kilkuset tysięcy). Niemniej jednak maszynowy odkrywca GELL-MANN postępuje w opisany sposób i udaje mu się osiągnąć sukces.

W moim przekonaniu wszystkie te okoliczności uzasadniają twierdzenie, że gdyby system GELL-MANN istniał i zostałby uruchomiony w roku 1964, to udało mu się naprawdę *odkryć*, a nie tylko *ponownie odkryć* kwarki. Odkrycie to dokonane byłoby jednak na poziomie praw fenomenologicznych i nadal brak byłoby głębszego rozumienia teoretycznego odkrytej przez system struktury, na poziomie ogólnych teorii wyjaśniających w rozumieniu Cartwright.

Należy stwierdzić, że jeśli „głównym ideologiem” komputerowych systemów odkryć był Herbert Simon, to „szarą eminencją” czy też „głównym wykonawcą” był Pat Langley, który jeszcze jako student dołączył do zespołu H. Simona i był autorem pierwszej wersji systemu BACON oraz współtwórcą wielu innych systemów, a po śmierci Herberta Simona i Jana Żytkowa nadal kontynuuje badania nad systemami odkryć. Langley wraz ze swoim zespołem obecnie zajmuje się głównie tzw. modelowaniem procesów indukcyjnych (Bridewell i Langley, 2010; Park, Bridewell i Langley,

2010), nowym kierunkiem badań w dziedzinie interaktywnych systemów odkryć naukowych, którego celem jest tworzenie skomplikowanych modeli matematycznych zjawisk na podstawie danych i wiedzy tła. W przeglądowym artykule podsumowującym ostatnią dekadę osiągnięć systemów odkryć Dzeroski, Langley i Todorovski (2007) powołują się na przykłady zastosowania systemów odkryć w zdobywaniu nowej, wartościowej wiedzy naukowej w kilku typach działalności naukowej: tworzeniu taksonomii, formułowaniu praw empirycznych oraz generowaniu modeli. We wszystkich tych przypadkach podkreślają oni możliwość interakcji między systemem a badaczami z niego korzystającymi jako ważną i pożądaną przez badaczy cechę systemu, choć w początkowym okresie rozwoju systemów odkryć grupy Simona, przeciwnie, kładziono nacisk na autonomię systemów. Coraz wyraźniej zaznaczające się nastawienie na praktyczne wyniki oraz to, że z Langleyem współpracują też badacze ze Słowenii związani z tradycją *machine-learning*, wskazuje na fakt coraz większego zbliżania się do siebie tych programów badawczych.

Podsumowując rozważania tej sekcji, wypada stwierdzić, że zarzuty krytyczne głoszące, iż systemy grupy Simona są zdolne co najwyżej powtórzyć pewne łatwe fragmenty procesu odkrycia naukowego dokonanego wcześniej przez naukowców, ponieważ brak im ścisłej, logicznej podbudowy metodologicznej niezależnej od powierzchownej, psychologicznej analizy rozumowań ludzkich, nie są słuszne w odniesieniu do wszystkich systemów tej grupy. GELL-MANN (i inne systemy, których celem jest znalezienie „ukrytej” struktury materii czy poszukiwanie ilościowych modeli zjawisk) rozumie w sposób indukcyjny, mechanicznie generując modele kwarkowe, następnie stosuje reguły heurystyczne pozwalające ograniczyć obszar przeszukiwania i odrzucić wystarczająco wcześnie nieadekwatne modele w duchu metodologii Poppera. Tym samym, wbrew ogólnym zarzutom krytyków skierowanym przeciwko systemom grupy Simona, system posługuje się innymi, niezależnymi metodami i rozumuje na innym poziomie teorii niż naukowcy w analogicznej sytuacji problemowej. Co więcej, o praktycznej wartości tego typu systemów dla nauki może świadczyć fakt, że coraz trudniej zlokalizować publikacje na ich temat, ponieważ coraz częściej pojawiają się one w periodykach specjalistycznych z dziedziny nauk szczegółowych, a nie w czasopiśmie poświęconych sztucznej inteligencji czy filozofii nauki.<sup>6</sup>

### 3.3. Program badawczy HHNT

Indukcja jako metoda zdobywania nowej wiedzy jest głównym tematem wspólnej książki Hollanda, Holyoaka, Nisbetta, i Thagarda, *Induction: Processes of Inference Learning and Discovery* (1986), która zapoczątkowała program badawczy zwany od nazwisk autorów HHNT. Program ten nie może się poszczycić serią tak licznych i skomplikowanych systemów odkryć opartych na wspólnych założeniach metodologicznych, jak program grupy Simona, ani praktycznymi zastosowaniami,

<sup>6</sup> Patrz np. Giza, 2006, s. 48; Langley, 2000; Dzeroski, Langley i Todorovski, 2007.

jak tradycja Turinga, niemniej jednak pod pewnymi względami wydaje się bardziej obiecujący niż pozostałe tradycje badawcze, ponieważ podejmuje trudne problemy, którymi żadna z nich się nie zajmuje: zbadanie mechanizmów indukcyjnych związanych z rozwojem wiedzy naukowej, także tej o charakterze teoretycznym.

W epilogu książki (s. 342) autorzy wyjaśniają jej główny cel badawczy: stworzenie podstaw ogólnej, pragmatycznej teorii indukcji jako ogólnych ram analizy kognitywnej procesów wnioskowania i uczenia się, poczynając od powstawania odruchów u szczurów aż po odkrycie naukowe.

Autorzy przedstawiają pojęciową analizę funkcjonowania tzw. modelu mentalnego, stworzonego przez system kognitywny w celu reprezentowania środowiska, w którym się znajduje, przy czym przyznają się oni (s. 23 n.) do związków zarówno z klasycznymi systemami produkcji (*production systems*) stosowanymi przez grupę Simona, jak i modelami koneksjonistycznymi. W pierwszym, wstępnym rozdziale książki znajdujemy pewne podstawowe zasady, leżące u podstaw zarówno implementacji komputerowych, jak i analiz jakościowych badanych przez autorów zjawisk. Podstawowe zasady funkcjonowania modelu są następujące:

— Wiedza reprezentowana jest przez prawa warunkowe o postaci warunek—akcja (*condition—action rules*). Działanie tych praw polega na generowaniu tzw. wiadomości dla całego systemu.

— Prawa mogą reprezentować zarówno zależności synchroniczne związane z kategoriami obiektów, jak i diachroniczne, wyrażające przewidywania na temat zachowania się obiektów. Oba typy praw współdziałają przy rozwiązywaniu problemów.

— System posiada mechanizmy indukcyjne, dzięki którym prawa mogą się łączyć w większe struktury, tzw. klastry. Tworzą one domyślne hierarchie oraz wyposażone są w prawa mające za zadanie „radzić” sobie z wyjątkowymi sytuacjami bez modyfikowania całej reszty domyślnej struktury.

— Prawa, których poprzedniki są spełnione przez aktywne w danej chwili wiadomości, konkurują ze sobą o reprezentowanie bieżącej sytuacji oraz wpływ na działanie systemu. Mogą one być uruchamiane równoległe do siebie i udzielać sobie nawzajem wsparcia, determinując w ten sposób zachowanie się systemu.

— Rozumowanie indukcyjne znajduje w systemie zastosowanie przy modyfikacji istniejących praw oraz generowaniu nowych. Mechanizmy generujące nowe prawa podlegają ścisłym ograniczeniom tak, aby prawa te okazały się przydatne dla systemu. Procedury indukcyjne posługują się wiedzą na temat obiektów i zdarzeń oraz sposobu, w jaki się one zmieniają.

Z powyższego wynika, że ogólna zasada działania systemu bliska jest standardowym systemom produkcji, a więc opiera się na cyklicznym dopasowywaniu do wzorca oraz uruchamianiu praw o postaci warunek—akcja. Zasadnicza różnica polega jednak na tym, że w standardowych systemach produkcji w danym cyklu uruchamiane jest tylko jedno prawo, a w modelu HHNT równoległe może być uruchomionych wiele praw, których warunki spełnione są przez wiadomości aktywne w systemie. Prawa te mają różny stopień „wsparcia” ze strony środowiska bądź innych aktyw-

nych w danym cyklu praw, a ogłaszane przez nie wiadomości mają wpływ na szanse uruchomienia innych praw w kolejnym cyklu. Taki sposób działania systemu daje mu znacznie większe szanse efektywnego działania w realnych sytuacjach, gdy wiedza o środowisku jest z reguły niekompletna lub nie do końca uświadomiona.

W rozdziale książki traktującym o odkryciu naukowym, autorzy podejmują próbę zastosowania swego modelu indukcji do pewnych aspektów tego, jak powiadają, „[...] największego indukcyjnego osiągnięcia człowieka” (Holland *et al.*, 1986, s. 320). Mówiąc w skrócie, pokazują, w jaki sposób wprowadzone przez nich pojęcie i mechanizmy, takie jak model mentalny środowiska i jego domyślne hierarchie, uogólnienia, tworzenie pojęć czy analogia, pozwalają wyeksplikować istotne elementy rozwoju wiedzy naukowej. Stawiają między innymi istotne pytanie, w jaki sposób naukowcy mogą sformułować nową teorię i nowe pojęcia, skoro odwołują się one do obiektów, które, przynajmniej na danym etapie rozwoju nauki, są niedostępne obserwacji i jak tę drogę zrekonstruować? Odpowiedź, której udzielają z perspektywy swego stanowiska, jest następująca: „[...] analogia jest głównym środkiem służącym konstrukcji teorii, a mechanizm kombinacji pojęciowej podstawową metodą generowania pojęć teoretycznych” (s. 326). Jest ona poparta jakościową analizą struktury, rozwoju i oceny teorii naukowych w ramach koncepcji modelu mentalnego oraz symulacją odkrycia prostej teorii naukowej o charakterze jakościowym za pomocą programu komputerowego PI (*Processes of Inference*) działającego według zasad tego podejścia.

Od czasu publikacji wspomnianej książki autorzy nie wydali żadnej wspólnej pracy, niemniej badania ich zaowocowały wieloma publikacjami na bardziej szczegółowe tematy oraz implementacjami komputerowymi działającymi w ramach wspólnie nakreślonego podejścia.

Jak wspomniałem wyżej, pierwszy z tych systemów, PI, został krótko opisany w pracy z 1986 roku, gdzie autorzy przedstawiają jego proste zastosowanie do odkrycia teorii falowej dźwięku na poziomie jakościowym przez analogię do fal na wodzie. W kolejnej książce, zatytułowanej *Computational Philosophy of Science*, Thagard (1988) podaje już wyczerpujący opis działania systemu w zastosowaniu do różnych aspektów odkrycia naukowego, dostrzega on jednak również jego istotne ograniczenia: system jest w stanie formułować jedynie bardzo proste prawa o charakterze jakościowym, brak mu bowiem, w przeciwieństwie do systemów powstałych w ramach dwu pozostałych tradycji, heurystyk pozwalających na wykrycie zależności o charakterze ilościowym. W kolejnej książce, zatytułowanej *Conceptual Revolutions* (1992), Thagard analizuje zmiany schematów pojęciowych towarzyszące rewolucjom naukowym za pomocą systemu PI oraz kolejnego systemu, ECHO (*Explanatory Coherence by Harmony Organization*).

W nowszych publikacjach autorzy skupiają się na roli analogii w twórczym myśleniu. W książce *Mental Leaps: Analogy in Creative Thought* (1995), Holyoak i Thagard opisują kolejne dwa systemy: ACME (*Analogical Constraint Mapping Engine*), i ARCS (*Analog Retrieval by Constraint Satisfaction*) i ich zastosowania do

kilku problemów z dziedziny tzw. fizyki jakościowej, takich jak analogia między przepływem wody i przepływem prądu elektrycznego czy między ruchem planet i ruchem elektronów w atomie. Systemy te są w stanie ustalić na drodze indukcji jedynie prawa o charakterze jakościowym związane z procesem tzw. transferu analogicznego, niemniej sędzę, że można je uznać za pierwszy krok w kierunku komputerowej symulacji procesu rozumowania przez analogię w odkryciu naukowym.

W kolejnych latach Thagard (1998) i jego współpracownicy zwrócili uwagę na rolę analogii o charakterze obrazowym czy wizualnym w strukturze i rozwoju wiedzy naukowej oraz odkryciu naukowym. Pierwsze komputerowe symulacje w tym kierunku to system DIVA (*Dynamic Imagery for Visual Analogy*), stworzony przy współpracy specjalistów od trójwymiarowej grafiki komputerowej (Croft and Thagard, 2002).

Nowsze prace Thagarda, Holyoaka i współpracowników idą w kierunku szczegółowego zbadania, metodami wypracowanymi przez kognitywistykę, roli analogii (w tym także o charakterze wizualnym) i kombinacji pojęciowej w twórczej działalności w wielu dziedzinach nauki od biologii i medycyny aż po same nauki komputerowe (*computer science*) (patrz np.: Lee i Holyoak, 2008; Saunders i Thagard, 2005; Thagard, 2005; 2008; 2010). Jak dotąd jednak, wedle mojej wiedzy, dociekania te nie zaowocowały nowymi systemami komputerowymi.

Podsumowując, jeśli podejście badaczy pracujących w tradycji Turinga można określić jako „logiczne” a grupy Herberta Simona (przynajmniej w początkowym okresie) jako „psychologiczne”, to podejście grupy HHNT należałoby nazwać mianem kognitywistycznego i koneksjonistycznego. Już z tego samego wynika, że systemy tej tradycji badawczej muszą być znacznie bardziej autonomiczne, ale też i bardziej skomplikowane niż systemy powstałe w pozostałych dwu programach badawczych. Brak spektakularnych praktycznych sukcesów mierzonych liczbą implementowanych systemów zdolnych do dokonywania ważnych odkryć o charakterze teoretycznym jest, w moim przekonaniu, podyktowany ogromną skalą trudności, na które napotyka tak ambitny program badawczy, i nie oznacza, że program ten ma błędne założenia czy metodologię. Przeciwnie, jest to jedyny program zdolny być może wyeksplikować tajemniczy akt twórczego myślenia naukowego za pomocą indukcyjnych mechanizmów kombinacji pojęciowej, abdukcji i rozumowania przez analogię i zaimplementować te mechanizmy w systemach zdolnych dokonywać odkryć naukowych.

#### 4. KONKLUZJE

Indukcja baconowska wraz z pewnymi ideami Poppera dotyczącymi falsyfikacji i odrzucania hipotez stała się dzięki „inteligentnym” systemom komputerowym częścią metody naukowej. Jest obecna w regułach wnioskowania logicznego systemów *machine learning* tradycji Alana Turinga zdolnych odkryć nowe, choć o bardzo niskim poziomie ogólności, prawa o charakterze statystycznym opisujące konkretne zjawiska. W systemach powstałych w ramach programu badawczego grupy Herberta

Simona metoda indukcyjna jest, wbrew twierdzeniom krytyków tego programu, stosowana przez systemy odkrywające ukrytą strukturę materii czy poszukujące modeli zjawisk. Systemy te używają innych metod i rozumują na innym, znacznie niższym poziomie teorii, niż czynią to naukowcy w danej dziedzinie badań.

Program badawczy HHNT nadal się rozwija, choć jak dotąd nie może poszczycić się sukcesami w postaci działających systemów zdolnych dokonywać rzeczywistych odkryć naukowych. Wynika to w dużej mierze ze skali trudności podjętych problemów. Celem jego jest bowiem analiza pojęciowa i implementacja komputerowa przy użyciu metod indukcyjnych, skomplikowanych procesów związanych z *autonomicznym* rozumowaniem systemu kognitywnego w tym także w przypadku dokonywania odkryć naukowych o charakterze teoretycznym. Zachęca to do optymizmu: być może w najbliższych latach będziemy świadkami powstania „inteligentnych”, stosujących metody indukcyjne autonomicznych systemów odkryć nowej generacji, choćby na początku miały one nawet bardzo wąski zakres zastosowań.

Na koniec pewna, być może nieco żartobliwa, uwaga. Otóż systemy stworzone w ramach tradycji Turinga czy przez badaczy z grupy Simona wydają się robić to, czego naukowcy od nich oczekują: nie starają się zastąpić i wyeliminować naukowców przez dążenie do względnej autonomii, lecz oferują im pomoc i współpracę. Patrząc z tej perspektywy, to, co filozofowie nauki i badacze sztucznej inteligencji uważają za niedostatek w programie badawczym HHNT, a więc brak praktycznych sukcesów, naukowcy być może powitaliby z niejaką ulgą. Ich kariery bowiem, jak twierdzi P. Langley (2002) w artykule poświęconym praktycznemu znaczeniu systemów odkryć dla nauki, obracają się właśnie wokół odkryć, których byli w stanie *samodzielnie* dokonać. Naukowcy potrzebują więc *pomocy* ze strony systemów komputerowych, a nie *automatycznych* (i *autonomicznych*) systemów odkryć.

## LITERATURA

- Bacon F. (1955), *Nowy Organon*, Warszawa: PWN.
- Bridewell W., Langley P. (2010), *Two Kinds of Knowledge in Scientific Discovery*, „Topics in Cognitive Science” 2, 36–52.
- Buchanan B., Feigenbaum E. (1978), *DENDRAL and meta-DENDRAL. Their Applications Dimension*, „Artificial Intelligence” 11, 5–24.
- Cartwright N. (1983), *How the Laws of Physics Lie*, Oxford: Oxford University Press.
- Croft D., Thagard P. (2002), *Dynamic Imagery. A Computational Model of Motion and Visual Analogy* [w:] *Model-Based Reasoning. Science, Technology, Values*, L. Magnani, N. Nersessian (red.), New York: Kluwer/Plenum, 259–274.
- Dzeroski S., Langley P., Todorovski L. (2007), *Computational Discovery of Scientific Knowledge* [w:] *Computational Discovery of Communicable Scientific Knowledge*, S. Dzeroski, L. Todorovski (red.), Berlin: Springer.
- Fischer P., Żytkow J. (1990), *Discovering Quarks and Hidden Structure* [w:] *Methodologies for Intelligent Systems*, Z. Ras, M. Zemankova, M. L. Emrich (red.), New York: Elsevier Science, 362–370.

- Gillies D. (1992), *Comments on „Scientific Discovery as Problem Solving” by Herbert A. Simon*, „International Studies in the Philosophy of Science” 6, 29–32.
- Gillies D. (1996), *Artificial Intelligence and Scientific Method*, New York: Oxford University Press.
- Giza P. (2002), *Automated Discovery Systems and Scientific Realism*, „Minds and Machines” 22, 105–117.
- Giza P. (2006), *Filozoficzne i metodologiczne aspekty komputerowych systemów odkryć naukowych*, Lublin: Wydawnictwo UMCS.
- Grobler A. (2006), *Metodologia nauk*, Kraków: Znak-Aureus.
- Holland J., Holyoak, K., Nisbett, R., Thagard, P. (1986), *Processes of Inference, Learning, and Discovery*, Cambridge (MA): MIT Press.
- Holyoak K., Thagard P. (1995), *Mental Leaps. Analogy in Creative Thought*. Cambridge: MIT Press.
- Langley P. (2000), *The Computational Support of Scientific Discovery*, „International Journal of Human-Computer Studies” 53, 393–410.
- Langley P. (2002), *Lessons for the Computational Discovery of Scientific Knowledge [w:] Proceedings of the First International Workshop on Data Mining, Sydney*, 9–12.
- Langley P., Simon H., Bradshaw G., Żytkow, J. (1987), *Scientific Discovery. Computational Explorations of the Creative Processes*, Cambridge (MA): MIT Press.
- Langley P., Żytkow J. (1989), *Data-Driven Approaches to Empirical Discovery*, „Artificial Intelligence” 40, 283–312.
- Lee H. S., Holyoak K. (2008) *The Role of Causal Models in Analogical Inference*, „Journal of Experimental Psychology: Learning, Memory, and Cognition” 34 (5), 1111–1122.
- Muggleton S., Feng C. (1992), *Efficient Induction of Logic Programs [w:] Inductive Logic Programming*, S. Muggleton (red.), London: Academic Press, 281–298.
- Muggleton S., King R., Sternberg M. (1992), *Protein Secondary Structure Prediction Using Logic-Based Machine Learning*, „Protein Engineering” 5/7, 647–657.
- Park C., Bridewell W., Langley P. (2010), *Integrated Systems for Inducing Spatio-Temporal Process Models*, „Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence”, Atlanta: AAAI Press.
- Popper K. (1961), *The Logic of Scientific Discovery*, New York: Science Editions (tłum. pol. Warszawa: PWN, 1977).
- Popper K. (1963), *Conjectures and Refutations. The Growth of Scientific Knowledge*, New York: Routledge & Kegan Paul (tłum. pol. Warszawa: PWN, 1999).
- Quinlan J. R. (1979), *Discovering Rules by Induction from Large Collections of Examples [w:] Expert Systems in the Microelectronic Age*, D. Michie (ed.), Edinburgh: Edinburgh University Press, 168–201.
- Quinlan J. R. (1986), *Induction of Decision Trees*, „Machine Learning” 1, 81–106.
- Rose D. (1988), *Using Domain Knowledge to Aid Scientific Theory Revision [w:] Proceedings of the Fifth International Workshop on Machine Learning*, Ithaca (NY): Morgan Kaufmann, 272–277.
- Saunders D., Thagard P. (2005), *Creativity in Computer Science*, [w:] *Creativity across domains: Faces of the muse*, J. C. Kaufman, J. Baer (red.), Mahwah (NJ): Lawrence Erlbaum Associates, 153–167.
- Simon H. (1992), *Scientific Discovery and Problem Solving*, „International Studies in the Philosophy of Science” 6, 1–14.
- Thagard P. (1988), *Computational Philosophy of Science*, Cambridge (MA): MIT Press.
- Thagard P. (1992), *Conceptual Revolutions*, Princeton: Princeton University Press.
- Thagard P. (1998), *Computation and the Philosophy of Science [w:] How Computers are Changing Philosophy*, T. Ward, J. Moor, (red.), New York: Blackwell, 48–61.

- Thagard P. (2005), *How to Be a Successful Scientist* [w:] *Scientific and Technological Thinking*, M. E. Gorman, R. D. Tweney, D. C. Gooding, A. P. Kincannon (red.), Mahwah (NJ): Lawrence Erlbaum Associates, 159–171.
- Thagard P. (2008), *Conceptual Change in the History of Science. Life, Mind, and Disease* [w:] *International Handbook of Research on Conceptual Change*, S. Vosniadou (red.), London: Routledge, 374–387.
- Thagard P. (2010), *Evolution, Creation, and the Philosophy of Science* [w:] *Epistemology and Science Education. Understanding the Evolution vs. Intelligent Design Controversy*, R. Taylor, M. Ferrari (red.), Milton Park: Routledge, 20–37.