

Jan Żytkow

Automatyzacja odkrycia naukowego: stan i perspektywy

Filozofia Nauki 1/4, 37-54

1993

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Jan Żytkow

Automatyzacja odkrycia naukowego: stan i perspektywy

Wprowadzenie

Teoria odkryć maszynowych to nowy, dynamicznie rozwijający się dział teorii sztucznej inteligencji, zajmujący się budową komputerowych systemów, które dokonują odkryć. Docelowo — automatyczny system odkryć można sobie wyobrazić jako podobny do człowieka-odkrywcy lub społeczności odkrywcow-naukowców robot, który przeprowadza eksperymenty i rozwija teorie na podstawie analizy uzyskiwanych danych.

Na odkrycia maszynowe można spojrzeć z dwu perspektyw: jako na modelowanie odkryć dokonywanych historycznie przez ludzi i jako na budowę normatywnych mechanizmów dokonujących odkryć. W pierwszym wypadku ważna jest zgodność funkcjonowania systemu z faktycznym przebiegiem procesu odkrycia, w drugim zaś ważna jest skuteczność systemu, jego zakres zastosowań, wydajność, czy prostota konstrukcji. W praktyce, podejście normatywne i modelowanie ludzkich odkryć wspierają się wzajemnie. Systemy normatywne są często oparte na analizach wypadków historycznych, które następnie uogólniają do postaci metod funkcjonujących w znacznie ogólniejszym zakresie. Stają się wtedy z kolei dobrym punktem wyjścia do tworzenia ogólniejszych modeli poznawczych.

W praktyce silniej reprezentowany jest program normatywny, choć przy analizie współczesnej nauki trudno jest czasem odróżnić podejście normatywne i poznawcze. Dla przykładu, analizujący dane naukowiec stosuje często metodę najmniejszych kwadratów, na ogół posługując się gotowym pakietem statystycznym. Rekonstrukcja tego pakietu jest więc niezbędna dla analizy poznawczej, która powinna być w stanie przewidzieć równanie uzyskane przez naukowca. Zrekonstruowany algorytm może stać się jednocześnie fragmentem systemu normatywnego. Rozróżnienie między podejściem normatywnym i poznawczym pojawia się, gdy pytamy o dalsze szczegóły. Na przykład, pytania, według jakich strategii naukowiec przeszukuje przestrzeń modeli (równań) i jak wybiera najlepszy model spośród modeli uzyskanych za pomocą

analizy statystycznej, są pytaniami zasadniczymi przy analizie ludzkich schematów odkryć, a więc przy podejściu poznawczym. Natomiast przy podejściu normatywnym ważna jest wydajna kontrola procesu przeszukiwania, unikająca wielokrotnej analizy tego samego modelu, a jednocześnie zapewniająca jak najobszerniejszy zakres przeszukiwania.

Związek odkryć maszynowych z historią i filozofią nauki może sięgać bardzo głęboko. Podejście poznawcze jest w gruncie rzeczy komputerowo uprawianą historią nauki, natomiast podejście normatywne — komputerową filozofią nauki. W obu wypadkach rekonstrukcja komputerowa procesu odkrycia zmusza do dbałości o szczegóły i prowadzi do znacznie pełniejszego modelu odkrycia niż jest możliwe przy użyciu środków tradycyjnych. Elementy pominięte w rekonstrukcji komputerowej ujawniają się automatycznie przy próbach zastosowania systemu komputerowego. Choć można wątpić, czy kiedykolwiek będzie zbudowany system, który modeluje procesy poznawcze we wszystkich szczegółach, to jednak dążenie do kompletności jest podstawowym elementem programu rekonstrukcji komputerowej. Kompletności można oczekiwać na poziomie ważniejszych celów poznawczych, zaś zakres uzyskanej kompletności jest ważnym elementem oceny skuteczności modelu.

Normatywne systemy pokrywają nie tylko obszar filozofii nauki, lecz także logiki — rozumianej szeroko jako metoda zdobywania wiedzy — wzbogacając te dziedziny o nowe problemy i rozwiązania. Reguły wnioskowania są uzupełniane przez mechanizmy kontrolne, sterujące procesem odkrycia. Kolejność stosowania reguł wnioskowania, sposoby podstawiania konkretnych wartości za zmienne w regule, metody selekcji wyników, czy złożoność obliczeniowa różnych metod są przykładami problemów nieodzownych dla metody komputerowej. Reprezentacja wiedzy, jej przybliżona prawdziwość, błąd empiryczny itp., muszą zostać zdefiniowane konstruktywnie i wystarczająco szczegółowo.

Tradycyjne rozróżnienie między kontekstem odkrycia i kontekstem uzasadniania umieszczało badania nad odkryciami poza filozofią nauki, w dziedzinie psychologii, socjologii i historii nauki, za niemożliwą uważając logiczną analizę i teorię odkrycia [Popper, 1961]. Przekonanie to upada w konfrontacji z istniejącymi systemami odkryć. Po pierwsze, stanowią one konstruktywny dowód istnienia normatywnych mechanizmów odkryć. Elementy metody naukowej wdrażane przez te systemy mogą być oceniane z punktu widzenia ich prawomocności. Systemy mogą być porównywane między sobą ze względu na większą lub mniejszą adekwatność empiryczną, zakres dokonywanych odkryć, czy prostotę generowanych teorii. Po drugie, systemy odkryć integrują odkrywanie i uzasadnianie w funkcjonalną całość [Żytkow i Simon, 1988], wskazując na doniosłość sprzężenia między odkryciem i uzasadnieniem.

Budowa systemów, które dokonują odkryć, wprowadza nowy element do filozofii i historii nauki, rewolucjonizujący obie dziedziny. Skonstruowany system — to nie tylko teoretyczna prezentacja metody, lecz przede wszystkim narzędzie, którego używać można przy dokonywaniu odkryć. Gotowe systemy, których można używać w eksperymentach nad odkryciami, dostarczają nieporównywalnie silnych bodźców do dalszego myślenia, a wkrótce powinny też dostarczać wyników wartościowych dla nauki. Budowa systemów odkryć opiera się na wypróbowanym schemacie, w którym

konstrukcja systemu jest wspomagana przez eksperymentalną weryfikację i analizę teoretyczną. Jest to konstruktywna metoda uprawiania historii i filozofii nauki, otwierająca nowe, fascynujące możliwości. Proces konstrukcji systemów odkryć zmusza do drobiazgowej, konstruktywnej analizy procesu odkrycia, otwierając nowe obszary pytań.

Naukowcy i filozofowie nauki przyglądają się osiągnięciom odkryć maszynowych ze sceptycyzmem. Z podobnym sceptycyzmem traktowano kiedyś traktor w porównaniu z koniem. Postawa ta na szczęście zmienia się już i można oczekiwać, że niedługo będziemy świadkami powszechnej akceptacji metody komputerowej w historii i filozofii nauki, a także w samej nauce. Rewolucja związana z komputerowymi systemami odkryć jest nieuchronna. Każdy może być jej uczestnikiem, włączając się do prac nad automatyzacją odkryć. Żeby ułatwić chętnym włączenie się do społeczności konstruktörów maszynowych odkrywców, dokonujemy niżej przeglądu istniejących systemów i szkicujemy wyłaniający się z ich analizy schemat teoretyczny. Ponieważ nie jest możliwe zrozumienie metod wypracowanych w dziedzinie odkryć maszynowych tylko na podstawie lektury artykułu przeglądowego, załączamy obszerny przegląd literatury pomocnej przy dalszych studiach.

Automatyzacja odkrycia

Teoria sztucznej inteligencji zajmuje się automatyzacją różnych czynności umysłowych wymagających inteligencji. Teoria odkryć maszynowych jest jednym z jej działów, stosunkowo jeszcze niewielkim, ale szybko rozwijającym się. W amerykańskiej społeczności pracującej nad sztuczną inteligencją dominuje przekonanie, że rozwój w tej dziedzinie następuje najefektywniej w cyklu budowy systemów komputerowych, empirycznej analizy ich funkcjonowania, wyciągania wniosków teoretycznych i znajdowania dalszych problemów, po którym następuje kolejny krok rozbudowy systemu. Prace w dziedzinie odkryć maszynowych prowadzone są w taki właśnie sposób.

Zazwyczaj bierze się jakąś umiejętność, którą posiadają odkrywcy w nauce, i konstruuje się system komputerowy, który w danych wejściowych odkrywa to, co odkryłby człowiek analizujący te same dane. Dobrym przykładem jest umiejętność znajdowania równań i innych formuł stanowiących indukcyjne uogólnienie danych wejściowych. Analiza funkcjonowania konkretnego systemu szybko wskazuje na różne jego ograniczenia, takie jak brak umiejętności znajdowania funkcji okresowych, niewłaściwe traktowanie błędu pomiaru itp. W kolejnych fazach rozwoju systemu istniejące ograniczenia są eliminowane, a raczej łagodzone, bo w sensie praktycznym żadna umiejętność nie jest nabywana w sposób absolutny, bez żadnych ograniczeń.

„Dane wejściowe” są pojęciem dwuznacznym. W szerszym sensie znaczą one wszystko to, co algorytm traktuje jako dane. W sensie węższym, ale zgodnym ze znaczeniem terminu w nauce, dane wejściowe to zbiór danych naukowych, na przykład wyników eksperymentów. W tym artykule rezerwujemy termin „dane” dla da-

nych w sensie naukowym, a dane wejściowe w sensie szerszym nazywać będziemy „sytuacją początkową”, bądź krócej „sytuacją”.

Do istniejącego systemu można kolejno dodawać nowe umiejętności, rozszerzając zakres czynności odkrywcy i prowadzące do nowych odkryć. Integrować można też gotowe systemy. Gdy różne umiejętności odkrywcy zostaną osobno zrekonstruowane jako systemy komputerowe, można tworzyć system, który łączy je razem. Na przykład, umiejętność znajdowania równań i umiejętność celowego gromadzenia danych na podstawie eksperymentów, mogą być połączone w systemie, który przeprowadza eksperymenty, gromadzi dane i znajduje indukcyjne ich uogólnienia w postaci równań [BACON: Langley, 1979; FAHRENHEIT: Żytkow 1987]

Zarówno składanie różnych systemów, jak i dodawanie nowych umiejętności do istniejącego systemu, to procesy złożone. Prosta integracja, polegająca na przekazywaniu wyjść jednego systemu jako wejść innych systemów, prowadzi do «kombinatorycznej eksplozji» złożoności obliczeniowej zintegrowanego systemu. Drugim poważnym problemem integracyjnym jest ograniczony zakres funkcjonowania każdego ze składników. Żadna umiejętność nie daje się praktycznie skomputeryzować w pełnym zakresie. Na przykład, system znajdujący równania może przeszukiwać nieograniczoną klasę funkcji, ale klasa ta na ogół zawiera tylko niewielki fragment zbioru wszelkich funkcji. Często dane empiryczne są lepiej wyjaśniane przez kilka wyrażeń funkcyjnych, z których każde obowiązuje w ograniczonym zakresie, zaś konkretny system znajdujący funkcje może w ogóle nie rozpatrywać tej możliwości. Inny przykład — to umiejętność odróżniania «złych» danych, będących rezultatem poważniejszego zakłócenia w procesie ich gromadzenia.

Integracja może być wykorzystana dla stworzenia bardziej skutecznego systemu. Na przykład, jeśli gromadzenie danych eksperymentalnych można «przeplatać» z analizą tych danych, wówczas znalezione w wyniku analizy równania lub granice zastosowań znalezionych równań mogą być wykorzystane przy planowaniu kolejnych serii eksperymentów [FAHRENHEIT: Żytkow i Zhu, 1993].

W pracy nad akumulacją możliwości odkrywczych nieodzowne są eksperymenty nad rozwijającym systemem. Jak wskazuje praktyka, sprzężenia zwrotne między częściami systemu ujawniają zaskakujące problemy i nasuwają idee nowych rozwiązań oraz lepszej integracji.

Podstawowe pojęcia przeszukiwania

W dziedzinie teorii sztucznej inteligencji akceptuje się powszechnie, że heurystyczne przeszukiwanie jest nieodzownym składnikiem inteligentnej aktywności [Simon, 1979; Nilsson, 1980]. Zdecydowana większość problemów nie daje się rozwiązać przez bezpośrednie algorytmy, prowadzące wprost do celu, bez eksploracji dróg, które niczego nie wnoszą do konstrukcji końcowego rozwiązania. Konieczne są tymczasowe kroki i ewaluacja ich przydatności, prowadzące do przeszukiwania poprzez próby i ocenę ich przydatności. Przestrzeń problemowa, nazywana też „przestrzenią przeszukiwania” i „przestrzenią stanów”, została wprowadzona w teorii sztucznej in-

teligencji jako abstrakcyjne narzędzie pojęciowe dla teoretycznego opisu procesu przeszukiwania [Simon, 1979; Nilsson, 1980]. Stany, operatory, ewaluacja, drzewa i grafy przeszukiwań, dostarczają teoretycznej podstawy dla rozwiązywania problemów za pomocą przeszukiwań.

Przestrzeń przeszukiwania składa się ze zbioru stanów S i z 2-argumentowej relacji E na S , nazywanej „relacją ekspansji”. E zawiera wszystkie bezpośrednie przejścia od stanu do stanu. W praktyce, stany nie są dane z góry, ale są konstruowane przez algorytm przeszukiwania ze stanów już istniejących. Przejście od stanu s_1 do stanu s_2 polega na zastosowaniu operatora, który na podstawie istniejącego stanu s_1 tworzy nowy stan s_2 . Prosty problem przeszukiwania można zdefiniować w drodze podania podzbioru S , zawierającego tak zwane stany początkowe, i innego podzbioru, zawierającego stany końcowe, wraz z zadaniem znalezienia trajektorii w przestrzeni stanów, prowadzącej od jakiegoś stanu początkowego do końcowego. Zwrotne i tranzytywne domknięcie E relacji E określa osiągalność stanów podczas wyczerpującego przeszukiwania.

Przeszukiwanie rozpoczyna się zazwyczaj od jednego stanu. Kolejne stany, będące węzłami w drzewie lub grafie przeszukiwania, są tworzone ze stanów istniejących przez operatory przeszukiwania. Operatory są algorytmami, wdrażającymi relacje E . Każdy operator może być inicjowany na wiele sposobów, prowadzących od tego samego stanu wejściowego do różnych stanów wyjściowych. Technicznie rzecz biorąc, różne inicjacje polegają na podstawianiu stałych za zmienne parametry występujące w operatorze. Wszystkie możliwe zastosowania operatora tworzą jego przestrzeń inicjacji. Operatory mogą być inicjowane za pomocą informacji znajdującej się w stanie wejściowym lub też za pomocą mechanizmu defaultów, jeśli stan wejściowy nie zawiera wystarczającej informacji lub też gdy nie wiadomo, jak informacje te wykorzystać. Brak informacji jest zastępowany przeszukiwaniem, w którym tworzy się wiele alternatywnych stanów, pomiędzy którymi dokonuje się później wyboru, gdy już wiadomo, jak wybór taki można przeprowadzić.

Skonstruowane stany są ewaluowane za pomocą różnych testów. Stan może być zaakceptowany lub odrzucony przez test boolowski, lub otrzymuje wartość numeryczną, która może być porównywana z wartościami przyporządkowywanymi innym stanom. Stany odrzucone przez ewaluatory nie są dalej rozwijane. Stany, które uzyskują niższe wartości numeryczne, są rozwijane w dalszej kolejności. W procesie przeszukiwania ewaluatory grają rolę przeciwną do operatorów, ograniczając przeszukiwanie, podczas gdy operatory rozszerzają przeszukiwanie do nowych stanów.

Kolejnym elementem przeszukiwania jest mechanizm kontrolny. Globalny mechanizm wybiera stan s z listy stanów aktualnie otwartych. Stan jest otwarty, gdy został skonstruowany, ale nie wszystkie zastosowania operatorów na tym stanie były już próbowane. Następnie lokalny mechanizm kontrolny powiązany z s wybiera jedno z możliwych zastosowań operatora do s , prowadzące do innego stanu. Mechanizm kontrolny reguluje kolejność tworzenia stanów. Typowe mechanizmy - to przeszukiwanie w głąb, wszerz, i przeszukiwanie, które wybiera najbardziej obiecujące stany.

Przeszukiwanie w głąb jest zwykle ograniczane przez maksymalną głębokość przeszukiwania. Robi się to dla uniknięcia nieskończonych ścieżek w drzewie przeszukiwania i dla przeprowadzenia wyczerpującego przeszukiwania do określonej głębokości. Przeszukiwanie wszędzie rozpatruje wszystkie prostsze stany przed stanami bardziej złożonymi, a więc łatwo je zatrzymać na najprostszym rozwiązaniu. Może ono jednak wymagać niedostępnie dużych pamięci, gdyż musi zapamiętywać znacznie większą liczbę stanów.

Przeszukiwanie typowe w systemach odkryć jest procesem stopniowej konstrukcji wiedzy. Operatory uzupełniają stan wiedzy o nowe dane, nowe pojęcia, nowe hipotezy itp., zaś ewaluatory oceniają zaproponowane konstrukty. Wiedza uzupełniana jest stopniowo, krok po kroku, przez kolejne zastosowania operatorów. Alternatywne wersje wiedzy są proponowane na alternatywnych gałęziach drzewa przeszukiwania.

Przegląd historyczny

Poniższy przegląd historyczny wskazuje główne kierunki badań nad automatyzacją odkryć i systemy, które kierunki te reprezentują, wraz z odnośnikami do literatury. Dopiero zapoznanie się ze źródłami opisującymi istniejące systemy i ich rezultaty pozwolić może na pełniejsze zrozumienie poszczególnych kierunków i ocenę całego programu.

Okolo roku 1980-ego dziedzina odkryć maszynowych była znana dzięki kilku systemom, głównie DENDRAL, BACON i AM [DENDRAL: Lindsay, Buchanan, Feigenbaum i Lederberg 1980; BACON: Langley 1979, Langley, Simon, Bradshaw i Żytkow 1987; AM: Lenat 1977, 1982]. Pracowało w niej nie więcej niż kilkanaście osób. Od tego czasu odkrycia maszynowe przechodzą wykładniczy wzrost, podwajając liczbę osób i publikacji co 3-4 lata. «Machine discovery» wyodrębnia się obecnie jako niezależny dział sztucznej inteligencji.

DENDRAL zajmował się komputerową rekonstrukcją struktury molekul związków organicznych, tworząc dla danego zestawu atomów grafy, opisujące wszystkie możliwe izomery zawierające te atomy. Liczba izomerów rośnie do nieosiągalnych praktycznie rozmiarów nawet dla niedużych ilości składników atomowych i przekracza poważnie liczbę izomerów faktycznie występujących w przyrodzie. Dlatego potrzebne są reguły, które narzucają ograniczenia na proponowane struktury. Meta-DENDRAL [Buchanan i Mitchell 1978] odkrywał, za pomocą danych ze spektroskopu masowego, takie reguły indukcyjne, które ujawniały fragmenty struktur molekul organicznych. DENDRAL i meta-DENDRAL były pierwszymi systemami w dziedzinie odkryć struktury, natomiast BACON znajdował równania empiryczne, które pasują do wejściowych danych empirycznych. Późniejsze wersje BACONa przeprowadzały eksperymenty, których wyniki uzyskiwane były za pomocą symulacji, uogólniły rekurencyjnie mechanizm planowania eksperymentów i znajdowania równań do większej liczby wymiarów, oraz wprowadziły mechanizm tworzący «intrinsic variables», tzn. zmienne liczbowe dla opisu zmiennych o charakterze symbolicznym [Langley i inni

1987]. Przykładami takich zmienych jest masa, różne ciepła właściwe i współczynniki przewodnictwa.

AM dokonywał odkryć głównie w indukcyjnie uprawianej arytmetyce liczb naturalnych i teorii zbiorów. Po przeprowadzeniu szeregu eksperymentów na obiektach matematycznych, AM wprowadzał nowe pojęcia i prawa potwierdzone przez te eksperymenty [Lenat 1977, 1982]. W porównaniu z BACONem, AM dysponował znacznie większą liczbą metod heurystycznych, to jest metod, które różnicowały zachowanie się systemu w różnych sytuacjach. Wychodząc od najprostszych pojęć teorii mnogości, AM wprowadził pojęcie „liczby”, arytmetyczne operacje na liczbach, pojęcie „liczby pierwszej” i zaproponował hipotezę Goldbacha, ale nie był w stanie zademonstrować innych zastosowań, co sugeruje, że system ten uchwycił niewiele z ogólnych mechanizmów odkrycia.

W pierwszej połowie lat 80-ych pojawił się szereg nowych systemów. GLAUBER [Langley i inni 1987] zastosował do danych jakościowych mechanizm wnioskowań indukcyjnych podobny do BACONA, uogólniając dane o relacjach do prawidliwości, które łatwo wyrazić w rachunku logicznym pierwszego rzędu — takich, jak to, że każda sól składa się z kwasu i zasady. GLAUBER grupował obiekty w klasy i tworzył nowe pojęcia przydatne do wypowiedzania prawidliwości.

STAHL [Żytkow i Simon 1986] analizował reakcje chemiczne, wyrażane w jakościowych terminach substancji na wejściu i wyjściu. Na ile pozwoliły na to dane, system ten stwierdzał, które z substancji są pierwiastkami, stwierdzał skład związków w terminach pierwiastków, oraz ukrytą strukturę reakcji chemicznych. STAHL potrafił identyfikować substancje występujące pod różnymi nazwami w różnych reakcjach.

DALTON [Langley i inni 1987] wykonywał kolejny krok w analizie ukrytej struktury. Rozpoczynając od wiedzy o pierwiastkach chemicznych i strukturze związków, uzyskiwanej przez system STAHL, oraz od danych na temat objętości substancji gazowych biorących udział w reakcjach, DALTON postulował skład atomowy molekul substancji chemicznych.

Liczba systemów odkryć i liczba kierunków rozwojowych wzrosła poważnie od połowy lat 80-ych, wzbogacając odkrycia maszynowe o wiele nowych umiejętności. Każda z wielowymiarowych przestrzeni empirycznych eksplorowana przez BACONa zawierała ograniczoną ilość wiedzy o analizowanych zjawiskach przyrodniczych. BACON odkrywał prawa empirycznie o postaci formuł matematycznych, ale nie analizował struktury tych formuł. Empiryczny kontekst praw został rozszerzony przez IDS [Nordhausen i Langley 1990], BLAGDEN [Sleeman, Stacey, Edwards i Gray 1989] i GALILEO [Żytkow 1990]. IDS i GALILEO potrafią reprezentować obiekty, stany i procesy oraz mogą rozumować na ich temat.

BACON dysponował bardzo prymitywnym rekurencyjnym mechanizmem projektowania eksperymentów. Zdolność ta została znacznie rozszerzona w systemach FAHRENHEIT [Żytkow 1987, Żytkow i Zhu 1993] i KEKADA [Kulkarni i Simon 1987]. KEKADA koncentrował się na projektowaniu doświadczeń, zmierzających do ulepszania substancji i sytuacji empirycznych tak, by odkrywane prawa miały szczególnie prostą formę. FAHRENHEIT modyfikuje swe strategie eksperymentowa-

nia pod wpływem odkrywanej wiedzy, poszukując granic zastosowań odkrytych praw i przeszukując nowe obszary, w których prawidłowości nie zostały jeszcze odkryte. Dwaj inni maszynowi odkrywcy, LIVE [Shen 1993] i DIDO [Scott i Markovitch 1993], używają innych mechanizmów selekcji eksperymentów. LIVE modeluje zachowanie ludzi, którzy nastawieni są nie na odkrycia, lecz konkretne problemy, i którzy dopiero w trakcie rozwiązywania postawionych problemów zmuszeni są do dokonywania odkryć. System ten podporządkowuje empiryczną eksplorację danej dziedziny i generację wiedzy o dziedzinie rozwiązywanym problemom, przeprowadzając obserwacje wtedy, gdy istniejąca wiedza nie pozwala na dokonywanie konkretnego przewidywania, bądź gdy przewidywanie okazało się fałszywe i wiedzę trzeba poprawić.

DIDO wnosi perspektywę probabilistyczną, wyrażając teorię jako sieć reguł probabilistycznych, przewidujących, z określonym prawdopodobieństwem, alternatywne następstwa tych samych sytuacji. DIDO ulepsza te reguły tak, by stały się bardziej deterministyczne, ogniskując nowe serie eksperymentów na obszarach, w których istniejąca wiedza jest najbardziej niepewna.

Ważna dla naukowca zdolność budowania nowych teorii przez analogię do teorii znanych była szeroko analizowana od połowy lat 80-ych. Zajmowali się tym Falkenhainer [1987], Falkenhainer i Rajamoney [1988], i wielu innych, zazwyczaj reprezentując teorie za pomocą tak zwanej jakościowej fizyki (*qualitative physics*: [Forbus 1984]).

Poważnych postępów dokonano w dziedzinie odkrywania i analizy równań empirycznych. Błąd pomiaru był używany przy ewaluacji hipotez już w najwcześniejszej wersji systemu BACON-1, ale zazwyczaj traktowano błąd w sposób bardzo uproszczony, zaniedbując propagację błędów do nowo tworzonych zmiennych, co często prowadzi do paradoksalnych rezultatów. Metody wyznaczania i stosowania błędów pomiaru używane we współczesnych naukach empirycznych zostały wprowadzone w systemie FAHRENHEIT i użyte przy odkryciach w laboratorium chemicznym [Żytków, Zhu i Hussam 1990]. Nauki empiryczne wymagają nie tylko umiejętności wyznaczania błędów pomiaru, lecz także jego redukcji do możliwie minimalnych rozmiarów. Żytków, Zhu i Zembowicz [1992, 1992a] wykazali, że metoda budowy teorii empirycznych, używana przez FAHRENHEIT, może być zastosowana do wykrycia teorii błędów pomiaru, a teoria błędów pomiaru może być zastosowana do analizy powtarzalności eksperymentów.

Generowanie równań z danych empirycznych jest jednym z najpopularniejszych zadań, wdrożonych w różnych wersjach. COPER [Kokar 1986] stosował metody analizy wymiarowej, ABACUS [Falkenhainer i Michalski 1986] znajdował równania stosujące się do części danych wejściowych i formuły, które określają zakres każdego z równań. KEPLER [Wu i Wang 1989], Equation Finder [Zembowicz i Żytków 1992] i E* [Schaffer 1990] — to przykłady innych systemów zajmujących się wynioskowywaniem równań z danych.

Naukowcy analizują równania empiryczne, wyciągając po różnych przekształceniach ważne wnioski. System GALILEO [Żytków 1990] prowadzi przeszukiwanie

w przestrzeni równoważnych form tego samego równania, przekształcając równania do postaci, która pozwala na ich uogólnienia.

Zapoczątkowana przez systemy DENDRAL, DALTON i STAHL rekonstrukcja metod odkrywania struktury materii rozwinięta została w różnych kierunkach. Wprowadzone przez STAHLa metody analizy reakcji zostały uogólnione przez systemy STAHLp [Rose i Langley 1986] i REVOLVER [Rose 1989] na reakcje między cząstkami elementarnymi. Zbudowany przez Fischera i Żytkowa [1990, 1992] system GELL-MANN stosuje się do różnorodnych odkryć struktury kwarkowej cząstek elementarnych, MECHEM [Valdes-Perez 1992, 1993] odkrywa strukturę reakcji chemicznych, zaś system MENDEL [Fischer i Żytkow 1992] dokonuje odkryć w dziedzinie genetyki. Sleeman i jego współpracownicy [1989] zasugerowali interesującą koncepcję przeszukiwania w przestrzeni jakościowych modeli układów chemicznych.

W miarę jak systemy odkryć potrafią reprezentować i odkrywać coraz bardziej skomplikowane teorie, rewizja odkrywanych teorii i weryfikacja metody odkryć przyciągają coraz więcej uwagi. Wprowadzona przez STAHLa metoda rewizji błędnych teorii została uogólniona przez STAHLp i REVOLVER na rewizje teorii cząstek elementarnych. Inną koncepcję rewizji teorii w dziedzinie cząstek elementarnych rozwinął Kocabas [1991], zaś dalsze metody rewizji teorii rozwinęli Rajamoney [1989, 1990], Shen [1993], Scott i Markovitch [1993].

Weryfikacja komputerowych metod odkrycia jest szczególnie zaawansowana w dziedzinie równań empirycznych. Schaffer [1990] porównał systematycznie równania empiryczne uzyskane przez BACON-1 i kilka innych systemów — z równaniami, otrzymanymi przez naukowców dla tych samych danych, wskazując na duże rozbieżności wyników. Zembowicz i Żytkow [1992] testowali na dużą skalę system Equation Finder za pomocą danych generowanych sztucznie ze znanych równań źródłowych z domieszką błędu, wykazując, że w miarę malenia błędu, odkrywane prawidłowości zbiegają do równań źródłowych.

W ostatnich latach znaczne zainteresowanie budzą zastosowania mechanizmów odkrycia do eksploracji baz danych w poszukiwaniu użytecznej wiedzy [Piatetsky-Shapiro i Frawley 1989; Piatetsky-Shapiro 1991; Żytkow 1992]. Dane dostępne w typowych bazach danych różnią się poważnie od danych gromadzonych przez naukowców eksperymentatorów [Żytkow i Baker 1991]. Dlatego systemy odkryć w bazach danych muszą używać innych technik przeszukiwania i reprezentacji wiedzy niż systemy odkryć naukowych.

Autonomia odkrywcy

Ponieważ we współczesnej nauce droga do poważnych odkryć wiedzie przez kombinacje wielu kroków, można je będzie dokonywać automatycznie dopiero wtedy, gdy nauczymy się, jak składać wiele kroków i metod. Istniejący maszynowi odkrywcy wykazują się sukcesami w powtórnym odkrywaniu znanych praw w stosunkowo prostych sytuacjach, choć mogą też oczywiście analizować zupełnie nowe dane. Czy system komputerowy, który odkrywa jakieś prawo znane historycznie,

może być nazwany „odkrywcą”? Wydaje się, że tak, gdyż o odkryciu X decyduje brak wiedzy o X przed odkryciem X , oraz brak wskazówek ze strony zewnętrznych autorytetów.

Uczenie się od zewnętrznego autorytetu (nauczyciela) jest znacznie łatwiejsze niż dokonywanie odkryć, ponieważ nauczyciel może kierować procesem uczenia na wiele sposobów. Na przykład, może przygotować eksperyment, o którym wie, że prowadzi do danych, które pozwolą na «odkrycie» pewnego równania empirycznego. Może określić zawnazu odpowiednią głębokość przeszukiwania, czy wybrać próg, od którego równanie będzie zaakceptowane. Nauczyciel może pomóc w selekcji rezultatów poszczególnych kroków odkrywczych. Komputerowe systemy uczące się korzystają z pomocy przy przygotowywaniu sytuacji wejściowej, przy ewaluacji częściowych rezultatów i przy kierowaniu poszczególnymi krokami. Ponieważ wszystkie szczegóły maszynowego odkrywcy dostępne są zewnętrznej inspekcji, daje to możliwość oceny rozmiarów zewnętrznej interwencji w dane odkrycie. W razie wątpliwości można w szczegółach przeanalizować wiedzę i metodę, które były punktem wyjścia do danego odkrycia i stwierdzić dokładnie, na ile odkrycie zostało dokonane samodzielnie.

Ludzie-odkrywcy, pojawiający się w dziejach nauki, nie opierali się na zewnętrznym autorytecie, ponieważ takiego jeszcze nie było w czasie, gdy dokonywali odkrycia, lub co gorsza, odkrycie zaprzeczało przekonaniom autorytetów. Zarówno indywidualny odkrywca, ludzkość jako zbiorowy odkrywca, jak i komputerowy system odkryć, muszą być wyposażone we własne, autonomicznie stosowane repertuary technik i wartości. Podczas gdy spolegliwy nauczyciel gwarantuje wartość przekazywanej wiedzy, odkrywca musi polegać na własnym osądzie.

Pojęcie autonomii wymaga komentarzy. Kepler, na przykład, odkrył swe prawa, posługując się danymi zebranymi przez Tycha de Brahe. Tak więc ani strategia gromadzenia danych, ani używane atrybuty nie pochodziły od niego. Jednak jego odkrycia były w dużej mierze autonomiczne, bo dane Tycha nie zawierały gwarancji, że ich analiza doprowadzi do odkryć; Kepler użył wybranych przez siebie hipotez, jak i wybranych przez siebie kryteriów sukcesu. Świadectwem jego początkowej niewiedzy jest duża liczba hipotez, którymi się posłużył, zanim odkrył swe prawa. Kryterium autonomii odkrycia stosuje się szczególnie dobrze do całego procesu historycznego, w mniejszym stopniu natomiast do poszczególnych epizodów historycznych. Wielkie odkrycia opierają się na wkładzie wielu osób w ciągu długiego czasu. Olbrzymia liczba obserwacji, pojęć i hipotez na temat ruchów planet była badana przed Keplerelem. Dopiero rozpatrując społeczność naukowców jako zbiorowego odkrywcę, widzimy cenę płaconą za uzyskanie wiedzy o właściwych atrybutach i danych, w efekcie których Kepler mógł dokonać swych odkryć.

Autonomię można zwiększać na wiele sposobów. Osoba działająca jest bardziej autonomiczna, gdy ma więcej środków do dyspozycji, na przykład, więcej sensorów czy manipulatorów. W ramach tych samych środków natomiast jest bardziej autonomiczna, gdy może dokonać więcej wyborów, zrealizować większą liczbę wartości i przebadać większy zakres celów.

Przyznać trzeba, że w zakresie autonomii dziedzina odkryć maszynowych nie wyprzedza znacznie innych dziedzin sztucznej inteligencji. Istniejące systemy odkryć

nie osiągnęłyby poważniejszych sukcesów w modelowaniu odkryć, gdybyśmy nie pomagali im z zewnątrz. Dla dalszego rozwoju dziedziny odkryć maszynowych kluczowe jest skoncentrowanie się na problemie autonomii i kolejnych kroków, prowadzących do jej zwiększenia przez redukcję zewnętrznej pomocy. Jeden ze sposobów zwiększania autonomii — to wdrażanie nowych składników procesu odkrycia. Niektóre składniki, takie jak mechanizmy konstrukcji nowych procedur pomiarowych [Żytkow, Zhu i Zembowicz 1992] i przyrządów, prowadzą do wynalazków raczej niż do odkryć, ale zaliczyć je należy do dziedziny odkryć maszynowych, gdyż poprzez nowe pomiary umożliwiają nowe odkrycia. Inne sposoby — to zwiększanie wewnętrznej integracji odkrywcy i autonomiczności dokonywanej ewaluacji. Większa autonomia oznacza więcej kroków dokonywanych kolejno bez zewnętrznej interwencji.

W sytuacji, gdy pewien rodzaj zewnętrznej interwencji zastąpiony jest przez automatyczne przeszukiwanie i gdy jednocześnie całkowite przeszukiwanie mieścić się musi w rozsądnych granicach, akumulacja kroków odkrywczych staje się poważnym wyzwaniem. Nacisk na akumulację jest jednocześnie poważnym bodźcem do stawiania nowych i ważnych problemów i uzyskiwania na nie doświadczalnie weryfikowalnej odpowiedzi. Pojedynczy krok rzadko daje wystarczającą perspektywę w ocenie rezultatów. Kiedy natomiast eliminujemy krok po kroku potrzebę zewnętrznej pomocy, jednocześnie utrzymując lub rozwijając zakres wiedzy, która może zostać odkryta, zbliżamy się do zrozumienia wartości naukowych i sposobów, na które są one powiązane.

Rozważmy, na przykład, dopasowywanie równań empirycznych do danych. Równania mogą być oceniane przez ich dopasowanie do danych, ale nawet gdy ograniczamy zakres dopasowywania do najprostszych równań, często kilka równań o porównywalnej prostocie pasuje do tych samych danych z podobną dokładnością. Każde z tych równań jest równie bliskim przybliżeniem do prawdy — na ile to można określić za pomocą danych wejściowych [Schaffer 1990; Zembowicz i Żytkow 1992]. Odkrywca może nie umieć lub nie chce dokonać wyboru, gdyż wybór ten może nie być trafny na dłuższą metę.

Sytuacja zmienia się, gdy większa autonomia odkrywcy daje mu większą perspektywę na dokonywane wybory, dostarczając dalszych kryteriów ewaluacji. Na przykład, może on zgromadzić dodatkowe dane w obszarze, w którym różne równania dają odróżnialne przewidywania. Dodatkowo — niektóre równania są bardziej podatne na generalizację, niektóre mogą redukować się do znanych teorii, niektóre pozwalają na lepszą interpretację parametrów. Wszystko to sprawia, że większa perspektywa uzyskana przez większy zakres czynności odkrywcy może pozwolić na dokonanie lepszego wyboru.

Jednym z ideałów zarówno w dziedzinie odkryć maszynowych, jak i w dziedzinie maszynowego uczenia się, jest autonomiczny, uczący się robot. Robot taki musi posiadać wiele umiejętności. Dopóki nie rozumiemy współdziałania różnych kroków w procesach odkrywania i uczenia się, małe są szanse, że zbudujemy uczącego się efektywnie robota. Od obecnych systemów do pełnej automatyzacji jest jeszcze długa droga. Rozwój w kierunku większej autonomii zadecyduje o przydatności, a więc o sukcesie maszynowych odkrywców.

Anatomia odkrywcy

Maszynowy odkrywca, którego koncepcje i techniczne rozwiązania szkicujemy poniżej, jest nadzbiorem różnych istniejących systemów, ale w swych możliwościach nie wybiega poza nie. Łączy on umiejętności robota do współoddziaływania ze światem zewnętrznym za pomocą sensorów i manipulatorów, różne algorytmy prowadzące do odkryć, reprezentację odkrytej wiedzy i pamięć o aktualnie wykonywanym układzie zadań odkrywczych.

Hardware odkrywcy musi zawierać komputer z jego procesorem, pamięcią i obsługą wejścia i wyjścia (I/O) oraz sensory i manipulatory połączone z I/O. Zapewniają one kontakt robota ze światem zewnętrznym [Żytkow, Zhu i Hussam 1990].

Software odkrywcy, niezbędny dla kontaktu ze światem zewnętrznym, musi zawierać programy kontrolujące poszczególne sensory i manipulatory (tzw. *device drivers*), oraz procedury operacyjne, które mogą wykorzystywać wiele elementarnych pomiarów i manipulacji, by uzyskać naukowo znaczące manipulacje i dane o świecie [Żytkow, Zhu i Zembowicz, 1992].

Wiedza systemu o świecie zawarta jest w sieci elementów, które reprezentują pojęcia, teorie, wiedzę o strukturze przeprowadzanego eksperymentu itp. [Nordhausen i Langley 1990; Żytkow i Zhu 1993; Rajamoney 1993]. Sieć ta rozwija się w miarę dokonywanych odkryć. Powiązanie wiedzy w sieć usprawnia dostęp do wiedzy, analizę stanu wiedzy i stawianie nowych celów.

Metoda odkrycia składa się z sieci zadań typowych w różnych warunkach, połączonych z planami, które mówią, jak zadania te można wykonać. Ponieważ zadania odkrywcy wymagają przeszukiwań, większość planów mówi o tym, jak dokonywać efektywnego przeszukiwania odpowiednich przestrzeni.

W konkretnej sytuacji, cele, plany i inne elementy metody odkrycia są konkretyzowane przez selekcję stałych, podstawianych za zmienne. Konkretnie zadania i konkretne plany działania zmieniają się dynamicznie, wzorując się na statycznej sieci typów zadań i planów. Dynamiczna — to znaczy przeprowadzana przez działający system — selekcja zadań i planów odbywa się nie tylko na podstawie wzorów czerpanych ze statycznej sieci typów, ale też na podstawie zgromadzonej wiedzy i na podstawie rozpatrywanych danych. Podobnie, konkretna wiedza reprezentowana jest w dynamicznej sieci rozwijanej na wzór statycznej sieci, reprezentującej typy wiedzy i ich powiązania [Żytkow 1991].

Statyczną sieć celów i planów oraz schematy reprezentacji wiedzy można traktować jako abstrakcyjnego odkrywcy. Konkretnych odkrywcy można tworzyć przez uzupełnienie abstrakcyjnego odkrywcy o różne zestawy sensorów i manipulatorów, wraz z programami kontrolującymi ich funkcjonowanie.

Cele odkrywcy

Dobrze zaprojektowany system odkryć składa się z szeregu oddzielnych modułów, przy czym każdy z nich funkcjonuje w określonym celu. Przegląd systemu na po-

ziomie zasadniczych celów pozwala na uproszczone i zrozumiałe jego przedstawienie. Każdy z celów musi być realizowany za pomocą konstruktywnych środków, które nazywać będziemy „planami”; zazwyczaj środki te oparte są na przeszukiwaniu. Ten sam cel może być realizowany za pomocą różnych planów. Na przykład wiele systemów (BACON, ABACUS, COPER, FAHRENHEIT, KEPLER, IDS) zawiera plany na znalezienie równania empirycznego pasującego do danych. Cele i plany można składać rekurencyjnie. Różne plany mogą stanowić alternatywy dla realizacji określonego celu, zaś realizacja określonego planu polega na osiągnięciu szeregu celów. Rekursja musi się ostatecznie odwoływać do planów, które są bezpośrednio wykonalne, bez odwoływania się do dalszych celów i planów.

Przeanalizujemy podstawowe elementy procesu odkrycia, w którym eksperymenty nad zaprojektowanym z góry układem fizycznym przeplatane są budową teorii. W typowym układzie eksperymentalnym można kontrolować szereg parametrów, tak zwanych zmiennych niezależnych, i dla każdego zestawu wartości tych zmiennych mierzyć odpowiedź układu fizycznego w terminach wartości zmiennych zależnych. Celem jest budowa teorii empirycznej, która opisuje, z dokładnością do błędu pomiaru, zależności między wybranymi zmiennymi niezależnymi i zależnymi.

Znajdywanie prawidłowości między jedną zmienną niezależną i jedną zmienną zależną jest ważnym celem w nauce. Prawidłowości takie są szczególnie proste i mogą być efektywnie uogólniane (BACON: [Langley i inni 1987]), a znajduwane być mogą za pomocą analizy danych, w których, przy ustalonych wartościach pozostałych zmiennych niezależnych, zmieniane są wartości jednej zmiennej niezależnej, i dla każdej z nich mierzona jest wartość zmiennej zależnej. Uzyskiwanie takich danych jest jednym z celów prowadzących do odkrycia równania empirycznego. Po przeprowadzeniu serii eksperymentów, ciąg wartości zmiennej niezależnej i odpowiadający mu ciąg wartości zmiennej zależnej są przekazane modułowi, który poszukuje równań pasujących do tych danych. Równanie takie może być znalezione bądź nie, co prowadzi w obu wypadkach do różnych nowych celów.

Po znalezieniu równania, alternatywne cele — to znalezienie granic jego stosowalności bądź uogólnienie równania do nowej zmiennej niezależnej (FAHRENHEIT: [Żytkow 1987]). Znalezienie granic stosowalności równania, to znaczy wartości zmiennej niezależnej, przy których równanie przestaje być spełnione, prowadzi do pytania o prawidłowości poza tą granicą. Cel ten jest realizowany w ten sam sposób, co znajdowanie pierwszej prawidłowości. Uogólnianie równania może odbywać się za pomocą rekurencyjnie wywoływanych celów gromadzenia danych empirycznych i znajdowania równań dla tych danych (BACON.3, FAHRENHEIT) oraz identyfikacji tych samych równań i innych obiektów, takich jak maksyma czy nieciągłości, znajdujących w różnych obszarach danych.

Jeśli nie udaje się znaleźć równania, pasującego do serii danych empirycznych, można dzielić te dane na krótsze serie, i stawiać zadanie znalezienia równania dla każdego fragmentu danych z osobna [Żytkow, Zhu i Zembowicz 1992, 1992a]. Podrzednym celem jest tu sensowna segmentacja danych, którą można przeprowadzać na podstawie maksymów, minimów, nieciągłości i tym podobnych specjalnych punktów, znajdujących w danych.

Omówiony zestaw celów wystarcza do budowy teorii empirycznej w N -wymiarowej przestrzeni wyznaczonej przez N zmiennych niezależnych, będących pod kontrolą eksperymentatora. Zanim jednak rozpocznie się budowa zasadniczej teorii, powinno się znaleźć błąd pomiaru i ulepszyć procedury operacyjne tak, by zredukować maksymalnie błąd pomiaru. Błąd pomiaru jest używany przy realizacji wielu celów, na przykład przy projektowaniu eksperymentów nad danym układem fizycznym, znajdowaniu równań i granic ich stosowności. Redukcja błędów prowadzi do dokładniejszych, bardziej powtarzalnych danych i w konsekwencji do wykrycia bardziej adekwatnych teorii. Zarówno określenie wielkości błędów jak i jego redukcję przez ulepszenie definicji operacyjnych dla mierzonych i kontrolowanych wielkości, da się przeprowadzić za pomocą tych samych środków, co budowę zasadniczej teorii [Żytkow, Zhu i Zembowicz 1992, 1992a]. Widać tu wyraźnie, że ten sam cel podrzędny, taki jak znalezienie równania empirycznego, może służyć do realizacji wielu celów nadrzędnych. W sumie liczba różnych celów i planów ich realizacji, niezbędnych dla szerokiego zakresu zastosowań systemu odkryć, może być nieduża.

Jeśli się wychodzi poza znalezienie układu równań empirycznych w przestrzeni N zmiennych niezależnych nad z góry przygotowanym układem fizycznym, to można iść w wielu kierunkach. Jeden z celów — to interpretacja znalezionych równań empirycznych, dzięki której nabierają fizycznego znaczenia wyrażenia składowe, takie jak energia kinetyczna poszczególnego obiektu. Transformacjami równań do postaci nadającej się do takiej interpretacji zajmuje się system GALILEO [Żytkow 1990]. Inny cel, który nabiera wagi w sytuacji konkurencyjnych teorii, między którymi nie da się rozróżnić za pomocą eksperymentów na danym układzie fizycznym, to przebudowa tego układu do postaci nadającej się do eksperymentów rozstrzygających między tymi teoriami [Rajamoney 1993].

Rola automatyzacji odkryć w filozofii nauki

Program automatyzacji odkryć prowadzi do nowej perspektywy na filozofię nauki. Filozofię nauki można traktować jako budowę maszynowych odkrywców i teorii ich funkcjonowania. Otwiera to nowe możliwości, nieporównywalnie bardziej atrakcyjne niż tradycyjna filozofia nauki. Maszynowi odkrywcy to systemy, które odtwarzają różne fragmenty metody naukowej. Można badać empirycznie ich skuteczność w konkretnych sytuacjach, można też analizować teoretycznie ich zakres zastosowań. Wszystkie szczegóły ich konstrukcji są dostępne analizie. Można zastępować różne konkretne rozwiązania przez inne i badać wpływ takiego zastąpienia na funkcjonowanie systemu. Kombinacja podejścia empirycznego i teoretycznego pozwala na szybki postęp i na prawdziwie naukowe podejście do filozofii nauki.

Kiedy się definiuje program filozofii nauki, to tradycyjnie umieszcza się filozofię nauki w ramach kontekstu uzasadnienia, rozumianego jako teoria norm naukowej weryfikacji. Maszynowi odkrywcy mieszczą się w normatywnym podejściu do odkrycia, rozszerzając zakres filozofii nauki. Automatyzacja odkryć pozwala widzieć filozofię nauki jako dziedzinę o wielkim znaczeniu praktycznym. Maszynowy od-

krywca to o wiele więcej niż teoria. To system, który może działać samodzielnie, dokonując odkryć, a przynajmniej wspomagając naukowca w najbardziej pracowitych i rutynowych elementach działalności odkrywczej.

Automatyzacja odkryć to program w filozofii nauki dający większe możliwości niż podejście tradycyjne. Jest to program konstruktywny i zmuszający do szczegółowej rekonstrukcji metody. W programie tym akumuluje się zarówno metoda, jak i odkrywana wiedza. Gotowe systemy dostarczają nie tylko lepszego zrozumienia wiedzy i metody naukowej, lecz także mogą być stosowane dla dokonywania odkryć.

Wiedza a metoda

Maszynowy odkrywca jest szczególnym źródłem wiedzy. Z pewnym uproszczeniem możemy porównać różne źródła wiedzy, takie jak książki, bazy danych, nauczyciele i odkrywcy, w dwuwymiarowej przestrzeni wiedzy o świecie i metody zdobywania tej wiedzy. System odkryć taki, jak IDS czy FAHRENHEIT, wyposażony jest w wyrafinowaną metodę, natomiast początkowa wiedza o świecie może być bardzo mała. Wiedza ta narasta w miarę dokonywanych odkryć, natomiast metoda stosowana przez system się nie zmienia. System może zachowywać się inaczej w obliczu nowych danych, ale tylko w ramach zawczasu zaprogramowanych rozwiązań. Systemy odkryć mogą posługiwać się mniej lub bardziej skomplikowaną metodą, ale metody tej same nie rozszerzają.

Dla kontrastu, trajektorie ilustrujące proces, który przechodzi każdy z nas, rozpoczynają się od małej ilości wiedzy i prostej metody. Potem zarówno nasza wiedza, jak i metoda się rozwijają, choć znaczną część naszej metody uczymy się od zewnętrznych autorytetów. Ludzkość, jako zbiorowy odkrywca, przechodzi podobną ewolucję, choć w jeszcze bardziej wyraźny sposób. Zarówno wiedza, jak i metoda rozwijały się w ciągu tysięcy lat od form prostych, ograniczonych i zawodnych — do wyrafinowanych i skutecznych metod i teorii współczesnej nauki. Tak więc maszynowy odkrywca różni się w swych możliwościach od odkrywców naturalnych.

Powstaje więc problem, czy można zbudować system, który będzie doskonalił metodę w miarę zdobywanej wiedzy. Przykłady znaczących elementów metody — to reprezentacja wiedzy przez równania, czy naukowa koncepcja błędów pomiaru. Pierwsza z nich stopniowo doskonaliła się poprzez średniowiecze i czasy nowożytne, obejmując formę równań, koncepcję zmiennych, transformację równań itp. O ile dla Arystotelesa mnożenie czy dzielenie dwu wielkości fizycznych nie miało sensu, przez co proste prawa mechaniki musiał formułować opisowo, o tyle w średniowieczu operacje te nabrały sensu, dając podstawy do rozwoju technik opartych na równaniach i ich transformacjach.

Pojęcie błędów pomiaru, techniki jego wyznaczania, propagacji i zastosowania w odkryciach i rozstrzyganiu sporów, rozwinęły się w wieku XIX-ym, a nabrały podstawowego znaczenia w wieku XX-ym. Lavoisier i jemu współczesni chemicy rozumieli, że pomiary nie są dokładne, ale nie rozumieli jeszcze zasad operowania błędem. Lavoisier, na przykład, dla udowodnienia zachowania masy w reakcjach che-

micznych wykazywał zachowanie masy wodoru i tlenu w reakcji tworzenia wody z wielką dokładnością, ale jego empiryczna metoda dawała proporcje tlenu do wodoru równą 5.7, znacznie odbiegającą od poprawnej wielkości 8.0.

Systemy takie jak IDS i FAHRENHEIT używają równań, a FAHRENHEIT używa nowoczesnej koncepcji błędu pomiaru. Czy da się w przyszłości zbudować system, który będzie rozbudowywał metodę? Generalnie odpowiedź musi być pozytywna, ale szczegóły są jeszcze bardzo niejasne. Jakie przestrzenie będą przeszukiwane podczas udoskonalania metody, jakie metody ewaluacji, jakie reprezentacje formalne? Jeśli system, który udoskonala metodę ma być modelem poznawczym dla rozwoju metody stosowanej przez ludzi, to jako punkt wyjścia należy przyjąć nie wyrafinowaną koncepcję przestrzeni alternatywnych metod czy ich składników, lecz raczej zestaw prostych pytań i prostych generatorów, przez których kombinacje wykształcić można wyrafinowane metody. Nie ma żadnej pewności, czy taki poznawczo adekwatny model zostanie kiedyś zbudowany.

Bibliografia

Buchanan, B. G., Mitchell, T. M. 1978. Model-directed learning of production rules, [w:] D. A. Waterman i F. Hayes-Roth (eds.), *Pattern-Directed Inference Systems*, New York: Academic Press.

Falkenhainer, B. C. 1987. Scientific Theory Formation Through Analogical Inference, *Proceedings of Fourth International Workshop on Machine Learning*, Los Altos, CA: Morgan Kaufmann Publ., s. 218-229.

Falkenhainer, B. C., Michalski, R. S. 1986. Integrating Quantitative and Qualitative Discovery: The ABACUS System, *Machine Learning* 1, s. 367-401.

Falkenhainer, B. C., Rajamoney, S. 1988. The Interdependencies of Theory Formation, Revision, and Experimentation, *Proceedings of the Fifth International Conference on Machine Learning*, Los Altos, CA: Morgan Kaufmann Publishers, s. 353-366.

Fischer, P., Żytkow, J. M. 1992. Discovering Quarks and Hidden Structure, [w:] Z. Ras, M. Zemankova i M. Emrich (eds.), *Proceedings of the Fifth International Symposium on Methodologies for Intelligent Systems (1990)*, New York: Elsevier, s. 362-370.

Fischer, P., Żytkow, J. 1992. Incremental Generation and Exploration of Hidden Structure, [w:] J. Żytkow (ed.), *Proceedings of the ML-92 Workshop on Machine Discovery*, Wichita, KS: National Institute for Aviation Research, s. 103-110.

Forbus, K. D. 1984. Qualitative Process Theory, [w:] D. G. Bobrow (ed.), *Qualitative Reasoning about Physical Systems*, Cambridge, MA: MIT Press.

Kocabas, S. 1991. Conflict Resolution as Discovery in Particle Physics. *Machine Learning* 6, s. 277-309.

Kokar, M. M. 1986. Determining Arguments of Invariant Functional Descriptions. *Machine Learning* 1, s. 403-422.

Kulkarni, D., Simon, H. A. 1988. The Processes of Scientific Discovery: The Strategy of Experimentation, *Cognitive Science* 12, s. 139-175.

Langley, P. W. 1979. Rediscovering Physics with BACON 3, *Proceedings of the International Joint Conference on Artificial Intelligence*.

Langley, P. W., Simon, H. A., Bradshaw, G., Żytkow J. M. 1987. *Scientific Discovery; An Account of the Creative Processes*, Boston, MA: MIT Press.

Lenat, D. B. 1977. Automated Theory Formation in Mathematics, *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, s. 833-842.

Lenat, D. B. 1982. AM: Discovery in Mathematics as Heuristic Search, [w:] R. Davis i D. B. Lenat (eds.), *Knowledge-based Systems in Artificial Intelligence*, New York: McGraw Hill, N.Y.

Lindsay, R., Buchanan, B. G., Feigenbaum, E. A., Lederberg, R. 1980. *Applications of Artificial Intelligence for Organic Chemistry; The DENDRAL Project*, New York: McGraw-Hill.

Nilsson, N. 1980. *Principles of Artificial Intelligence*, Palo Alto, CA: Tioga.

Nordhausen, B., i Langley, P. 1990. An Integrated Approach to Empirical Discovery, [w:] J. Shragar i P. Langley (eds), *Computational Models of Scientific Discovery and Theory Formation*, San Mateo, CA: Morgan Kaufmann Publishers, s. 97-128.

Piatetsky-Shapiro, G. (ed.) 1991. *Proceedings of AAAI-91 Workshop on Knowledge Discovery in Databases*, San Diego, CA.

Piatetsky-Shapiro, G., Frawley, W. (eds.) 1991. *Knowledge Discovery in Databases*, Menlo Park, CA: AAAI Press.

Popper, K. R. 1961. *The Logic of Scientific Discovery*. New York: Science Editions.

Rajamoney, S. 1989. *Explanation-Based Theory Revision: An Approach to the Problems of Incomplete and Incorrect Theories*, PhD thesis, Department of Computer Science, University of Illinois at Urbana-Champaign.

Rajamoney, S. 1990. A Computational Approach to Theory Revision, (w:) J. Shragar i P. Langley (eds.), *Computational Models of Scientific Discovery and Theory Formation*, San Mateo, CA: Morgan Kaufmann Publishers, s. 225-254.

Rajamoney, S. 1993. The design of Discrimination Experiments, *Machine Learning* 12, s. 185-203.

Rose, D. 1989. Using Domain Knowledge to Aid Scientific Theory Revision, *Proceedings of the Sixth International Workshop on Machine Learning*, San Mateo, CA: Morgan Kaufmann Publishers.

Rose, D., Langley, P. 1986. Chemical Discovery as Belief Revision. *Machine Learning* 1, s. 423-451.

Schaffer, C. 1990. A Proven Domain-Independent Scientific Function-Finding Algorithm, *Proceedings of the AAAI-90*, AAAI Press, s. 889-894.

Scott, P. D., Markovitch, S. 1993. Experience Selection and Problem Choice in an Exploratory Learning System, *Machine Learning* 12, s. 49-68.

Shen, W. M. 1993. Discovery as Autonomous Learning from Environment, *Machine Learning* 12, s. 143-166.

Simon, H. A. 1979. *Models of Thought*, New Haven, CON: Yale University Press.

Sleeman, D. H., Stacey, M. K., Edwards, P., Gray, N. A. B. 1989. An Architecture for Theory-Driven Scientific Discovery, *Proceedings of EWSL-89*.

Valdes-Perez, R. E. 1992. Theory-driven Discovery of Reaction Pathways in the ME-CHEM System, *Proceedings of National Conference on Artificial Intelligence*, s. 63-69.

Valdes-Perez, R. E. 1993. Conjecturing Hidden Entities via Simplicity and Conservation Laws: Machine Discovery in Chemistry, *Artificial Intelligence* [w druku].

Wu, Y., Wang, S. 1989. Discovering Knowledge from Observational Data, (w:) G. Piatetsky-Shapiro (ed.), *Knowledge Discovery in Databases, IJCAI-89 Workshop Proceedings*, Detroit, MI, s. 369-377.

Zembowicz R. i Żytkow, J.M. 1992. Discovery of Equations: Experimental Evaluation of Convergence, (w:) *Proceedings of the Tenth National Conference on Artificial Intelligence*, Menlo Park, CA: AAAI Press/MIT Press, s. 70-75.

Żytkow, J. M. 1987. Combining Many Searches in the FAHRENHEIT Discovery System, *Proceedings of the 4th International Workshop on Machine Learning*, Irvine, CA: Morgan Kaufmann, s. 281-287.

Żytkow, J. M. 1990. Deriving Basic Laws by Analysis of Processes and Equations, (w:) J. Shrager, P. Langley (eds.), *Computational Models of Discovery and Theory Formation*, San Mateo, CA: Morgan Kaufmann Publishers, CA. s. 129-156

Żytkow, J. M. 1991. Integration of Knowledge and Method in Real-world Discovery, *SIGART-91*.

Żytkow, J. M. (ed.) 1992. *Proceedings of the ML-92 Workshop on Machine Discovery (MD-92)*, Wichita, KS: National Institute for Aviation Research.

Żytkow, J. i Baker, J. 1991. Interactive Mining of Regularities in Databases, (w:) Piatetsky-Shapiro i W. Frawley (eds.), *Knowledge Discovery in Databases*, Menlo Park, CA: AAAI Press, s. 31—53.

Żytkow, J. M., Simon, H. A. 1986. A Theory of Historical Discovery: The Construction of Componential Models, *Machine Learning* 1, s. 107-136.

Żytkow, J. M. i Simon, H. A. 1988. Normative Systems of Discovery and Logic of Search, *Synthese* 74, s. 65-90.

Żytkow, J. M., Zhu, J. 1993. Experimentation Guided by a Knowledge Graph, (w:) Shen (ed.), *Proceedings of the AAAI-93 Workshop on Learning Action Models*, AAAI Press.

Żytkow, J. M., Zhu, J., Hussam, A. 1990. Automated Discovery in a Chemistry Laboratory, *Proceedings of the Eighth National Conference on Artificial Intelligence*, The AAAI Press, s. 889-894.

Żytkow, J. M., Zhu, J., Zembowicz, R. 1992. Operational Definition Refinement: a Discovery Process, *Proceedings of the Tenth National Conference on Artificial Intelligence*, AAAI Press, s. 76-81.

Żytkow, J. M., Zhu, J., Zembowicz, R. 1992a. The First Phase of Real-World Discovery: Determining Repeatability and Error of Experiments, *Machine Learning: Proceedings of the Ninth International Conference, July 1992, Aberdeen, United Kingdom*, Morgan Kaufmann Publishers, s. 480-485.