

**Janina A. Jakubczyc**

Uniwersytet Ekonomiczny we Wrocławiu

---

## KONTEKST W UCZENIU SIĘ POJĘĆ

---

**Streszczenie:** rosnące woluminy, wzrastająca złożoność, wielość źródeł, a także cele gromadzenia danych różne od realizowanych zadań analizy danych są jedną z przyczyn niepowodzeń algorytmów uczenia się pojęć. Szczególnego znaczenia nabierają upraszczające i nieadekwatne do rzeczywistości założenia. Dotyczą one precyzyjności pojęć, kontekstowej niezależności oraz możliwości reprezentacji pojęcia przez pojedynczy opis symboliczny. Przegląd sposobów uwolnienia tych założeń i poszerzenie opisu pojęć o zależności kontekstowe są przedmiotem artykułu.

**Słowa kluczowe:** maszynowe uczenie, uczenie się pojęć, kontekst.

### 1. Wstęp

Rosnące woluminy, wzrastająca złożoność, wielość źródeł, a także cele gromadzenia danych różne od realizowanych zadań analizy danych są jedną z przyczyn niepowodzeń algorytmów uczenia się pojęć. Szczególnego znaczenia nabierają upraszczające i nieadekwatne do rzeczywistości założenia. Dotyczą one precyzyjności pojęć, kontekstowej niezależności oraz możliwości reprezentacji pojęcia przez pojedynczy opis symboliczny.

Odnosząc się tylko do definicji pojęć naturalnych (punkt 2.1), należy koniecznie poszerzyć opis pojęć o zależności kontekstowe, które są ich nieodłączną częścią (punkty 2.2 i 2.3). Istniejące propozycje uczenia się elastycznych pojęć przyjmują również wiele założeń ograniczających ich rzeczywiste zastosowania. W zależności od charakteru dziedziny bądź trybu przetwarzania danych uczących, proponowane są różne metody identyfikacji kontekstu i sposoby wykorzystania informacji kontekstowej w procesie zarówno przygotowania danych, jak i uczenia (punkt 3) oraz formy jego reprezentacji (punkt 4). W artykule przedstawiono przegląd istniejących podejść i algorytmów oraz propozycje poszerzenia badań w zakresie uczenia się elastycznych pojęć (punkt 5).

## 2. Pojęcie, kontekst i rola kontekstu

### 2.1. Pojęcie

Algorytmy uczenia się pojęć determinowane są przyjętą definicją pojęcia. Pojęcie jest formą wiedzy o świecie służącą do opisu i interpretacji obserwacji lub abstrakcyjnych idei. Pojęcie jest opisem umożliwiającym identyfikację wspólnych cech pojęcia oraz cech odróżniających je od innych pojęć. W zależności od precyzyjności opisu rozróżnia się pojęcia matrycowe i naturalne [48; 49]. Pojęcie matrycowe to pojęcie dobrze określone. Stanowi ono poznawczą reprezentację skończonej liczby wspólnych cech, które w jednakowym stopniu przysługują wszystkim egzemplarzom. Istotną cechą desygnatów pojęcia matrycowego jest ich nierozróżnialność – stopień przynależności jest taki sam dla wszystkich egzemplarzy tego rodzaju pojęcia. Egzemplarz danego pojęcia matrycowego nie może być bardziej lub mniej typowy, lepszy lub gorszy, gdyż pojęcie takie traciłoby sens poznawczy. Reguła klasyfikacyjna przypisująca dany egzemplarz do danego pojęcia jest jednoznaczna. Tego rodzaju pojęcia są zdefiniowane precyzyjnie w obszarach matematyki, fizyki czy biologii.

W przeciwieństwie do pojęcia matrycowego, pojęcie naturalne, powszechne w życiu człowieka, charakteryzuje się możliwą niejednoznacznością i nieokreślonością. Jest ono reprezentacją poznawczą odzwierciedlającą zbiór wspólnych cech, które w dużym stopniu przynależą do egzemplarzy danego pojęcia. Pojęcie naturalne pozwala na rozróżnienie typowych i nietypowych przykładów pojęcia oraz przykładów lepszych i gorszych. Nieprecyzyjność pojęcia naturalnego oznacza, że granice pojęcia naturalnego zmieniają się w czasie lub z kontekstu na kontekst. Tworzenie reguły klasyfikacyjnej zależy od rozkładu jednostek w przestrzeni obserwowanych atrybutów; jedne obiekty mogą być bardziej typowe niż inne. Miary determinujące „typowość” z reguły opierają się na stopniu podobieństwa (według wybranych atrybutów opisujących) między danym obiektem a „średnim” obiektem w przestrzeni atrybutów. Wybór atrybutów opisujących zależy od ich wpływu na oryginalną zasadę klasyfikującą dane pojęcie, a ocena wpływu zależy od wiedzy i doświadczenia eksperta.

Takie same znaczeniowo rodzaje pojęć definiuje Michalski [33; 34], nazywając je dobrze zdefiniowanymi będącymi odpowiednikami pojęć matrycowych i niedokładnie zdefiniowanymi (elastycznymi), które są odpowiednikami pojęć naturalnych.

Bardziej szczegółowy podział pojęć proponują S. Matwin i M. Kubat [32], rozróżniając trzy rodzaje pojęć: absolutne, względne i częściowo względne. Pojęcie absolutne jest równoznaczne z pojęciem matrycowym. Zamiast pojęcia naturalnego wprowadzają dwa pojęcia: względne i częściowo względne. Pojęcie względne nie ma żadnej własności występującej we wszystkich okolicznościach. Natomiast pojęcia częściowo względne charakteryzują się zbiorem podstawowych atrybutów

zawsze obecnych w opisie pojęcia oraz zbioru cech zmieniającego się z kontekstu na kontekst. Ze względu na trudności, jakie mogą pojawić się w odróżnieniu pojęć względnych od częściowo względnych na początku procesu uczenia się pojęć, taki szczegółowy podział nie wydaje się zasadny.

Można powiedzieć, że utworzenie spójnego i jednoznacznego opisu pojęcia naturalnego z definicji nie jest możliwe. Jak zbliżyć się do jak najpełniejszego, najbardziej spójnego i najbardziej jednoznacznego opisu? Jednym z pomysłów jest uwzględnienie w opisie pojęcia wszystkich potencjalnych kontekstów, które umożliwią jego różne rozumienie i interpretację.

## 2.2. Kontekst

Przejdźmy zatem do zdefiniowania kontekstu. Jeśli chodzi o uniwersalność, na uwagę zasługuje kilka definicji. Dwie pierwsze podają V. Akman i M. Surav [1; 2]. Definicja przytoczona za słownikiem filozoficznym mówi, że kontekst jest całkowitą sumą znaczeń (skojarzeń, idei, założeń, przyjętych opinii), która jest blisko związana z daną rzeczą, zauważa jej pochodzenie, wpływa na postawę, perspektywę, oceny i wiedzę o rzeczy.

Natomiast definicja kontekstu ze słownika *Collins Cobuild English Language Dictionary* mówi o powszechnym znaczeniu tego terminu, które obejmuje idee, sytuacje, zdarzenia lub informacje powiązane z danym pojęciem, umożliwiające jego pełne zrozumienie. Widzenie jakiegoś pojęcia w kontekście oznacza jego rozważenie w relacji do wszystkich powiązanych z nim czynników.

Uogólniając te definicje, można powiedzieć, że kontekst pojęcia odnosi się do idei, sytuacji, zdarzeń lub informacji, które dają możliwość jego pełnego zrozumienia i interpretacji.

Dwie kolejne definicje kontekstu reprezentują spojrzenie na kontekst przez pryzmat jego efektów na dane pojęcie [17]. W tym przypadku kontekstem jest to wszystko, co powoduje zmiany w interpretacji i przypisywaniu znaczeń danemu pojęciu. G. Hirst porównuje kontekst do stresorów, które mają różny wpływ na różne rzeczy (osoby, przedmioty, sytuacje). Dlatego konteksty definiowane są względem ich efektów na dany system, rzecz, osobę.

Podobną w swej intuicyjności definicję podają P. Brézillon i J.-Ch. Pomerol [8], odnosząc ją do skutków, mówiąc, że kontekst to wymuszanie czegoś bez bezpośredniej interwencji („that which constrains something without intervening in it explicitly”).

Podsumowując, można powiedzieć, że kontekstem może być wszystko to, co pośrednio wymusza rozumienie i interpretację danego pojęcia, a więc pojedynczy czynnik lub sytuacja o różnym poziomie złożoności.

### 2.3. Rola kontekstu w uczeniu się pojęć

Możliwości, jakie daje zastosowanie kontekstu w uczeniu się pojęć, są nie do przecenienia. Po pierwsze, w każdym inteligentnym działaniu, a takim jest uczenie, niezbędne jest wspomaganie rozumienia nauczanych pojęć poprzez ich możliwe konteksty. Możliwość interpretacji złożonej inteligencji zawartej w opisach zjawisk, rzeczy i obiektów, jaką potencjalnie daje znajomość kontekstów, znakomicie powiększa zakres poznania i opisu rzeczywistości (np. opis pola walki [39]).

Kontekst może być sposobem na ograniczenie ogromnej ilości informacji i jednocześnie pełniejsze jej wykorzystanie [38]. Jest środkiem umożliwiającym koncentrację na istotnych aspektach danej sytuacji i ignorowanie innych aspektów, mogących zaburzać lub uniemożliwiać prawidłowe rozumienie i interpretację. Zdecydowanie zwiększa prawidłowy ogląd uczonych pojęć. Ponadto pełniejszy, bardziej wszechstronny opis daje szansę zwiększenia trafności klasyfikacyjnej wygenerowanych modeli pojęć [9; 37].

Zwiększenie zakresu zmienności interpretacji nauczanych pojęć decyduje o zwiększeniu możliwości poznania świata. Poszerzenie opisu o możliwe konteksty zwiększa z reguły liczbę atrybutów opisujących dane pojęcie, ale trafność klasyfikacyjna i precyzja opisu są większe. Ponadto nie ignoruje faktu, że powiązanie niektórych deskryptorów następuje tylko w danym kontekście. Umożliwia identyfikację charakterystyk, które są odpowiednie tylko w ograniczonej przestrzeni przykładów, a wysoce nieodpowiednie wszędzie indziej [12]. Kontekstowe podejście do opisu pojęć daje możliwość sformułowania szczegółowych kontekstów dla trudnych do rozwiązania problemów, np. w systemie CYC [29]. Pojawia się problem, skąd wziąć konteksty, w jakich może występować nauczane pojęcie?

## 3. Identyfikacja kontekstu

Niedoskonałość teorii próbujących wyjaśnić społeczno-ekonomiczną działalność człowieka i organizacji społeczno-gospodarczych determinuje spory margines niewiedzy o badanych zjawiskach. Stąd problem identyfikacji kontekstu może być istotnym elementem poznania nieznanych aspektów uczonych pojęć. Równie ważny w uczeniu się pojęć może być kontekst znany lub możliwy do identyfikacji przez ekspertów lub analityków.

Identyfikacja kontekstu determinowana jest przez jego lokalizację. Kontekst może się znajdować w rozważanym zbiorze opisującym dane pojęcie, ale również może być poza nim. Może występować pojedynczo lub grupowo w postaci wielu rozłącznych lub/i nachodzących na siebie kontekstów, a także pojawiać się jednocześnie lub cyklicznie.

Zadanie poszukiwania kontekstów zawartych w pliku danych opisujących dane pojęcie lub poza nim można postawić przed ekspertami z danej dziedziny, traktując

ich jako źródła wiedzy kontekstowej. Z reguły zatrudniani są, gdy próby utworzenia opisu pojęcia są nieudane. Wówczas dodatkowe informacje, którymi mogą posłużyć eksperci, wynikają z analizy wyników klasyfikacji i odniesienia ich do wiedzy dziedzinowej. Sugerowanym rozwiązaniem może być postulat uwzględnienia dodatkowych charakterystyk dotyczących bezpośrednio danego pojęcia, jak w zadaniu detekcji plam oleju [26]. Atrybuty wykorzystywane do tego zadania okazały się niewystarczające. Ekspert po analizie wyników zaproponowali poszerzenie zbioru atrybutów o atrybuty umożliwiające pogrupowanie obrazów według kontekstu i utworzenie oddzielnych opisów dla każdego kontekstu.

Innym sposobem jest identyfikacja atrybutu lub atrybutów kontekstowych w analizowanym zbiorze danych. Przykładem jest zadanie harmonogramowania spotkań uniwersyteckiej kadry naukowej, w którym analiza wyników klasyfikacji umożliwiła wykrycie okresów najgorszej predykcji [36]. Lokalizacja spotkań okazała się użytecznym kontekstem dla tego zadania.

Natomiast Y. Wang i in. [55] wyselekcjonowali dwie grupy cech kontekstowych: jedną zależną od struktury (miejsce wiadomości w strukturze), a drugą charakteryzującą sekwencyjność (miejsce umieszczenia wiadomości w sekwencji kodu) w uczeniu kolaboracyjnym.

Oprócz analizy wyników klasyfikacji, można poszukiwać kontekstu w zbiorach danych opisujących wiedzę dodatkową, jak pisze R.S. Michalski [34], gdy kontekst znajduje się poza zbiorem przykładów.

Nie zawsze możliwe jest wykorzystanie eksperta do identyfikacji możliwych kontekstów. W takich sytuacjach należy opracować jakiś algorytm automatycznie identyfikujący atrybuty kontekstowe, a następnie znaleźć sposób wykorzystania informacji kontekstowej. Koniecznym założeniem jest to, by informacje o różnych kontekstach były zawarte w danych uczących, a w przypadku zmienności pojęcia w czasie również w nauczonej klasyfikatorze, który należy modyfikować wraz ze zmianą kontekstu [15; 21; 51].

Na kontekst w uczeniu się pojęć patrzy się przez pryzmat zmienności uczonego pojęcia w czasie lub uczenia kontekstowego i zależności kontekstowej. Pierwsze realizowane jest w uczeniu przyrostowym (punkt 3.2), podkreślającym zmienność pojęć w czasie. Drugie jest rodzajem uczenia w trybie wsadowym (punkt 3.1).

### 3.1. Uczenie wsadowe

Uczenie kontekstowe i zależności kontekstowych możliwe jest, gdy konteksty zawarte są w danych. Z reguły jest to postępowanie dwufazowe. W pierwszym kroku identyfikuje się kontekst, a w drugim wykorzystuje się znaleziony kontekst w procesie przygotowania danych lub uczenia się pojęć. Jednym z proponowanych sposobów identyfikacji kontekstów jest rozróżnienie cech podstawowych (decyzyjnych, wykonawczych), niedecyzyjnych (nierелеwantnych) kontekstowo zależnych i

kontekstowych [30; 52; 53]. W niektórych pracach uszczegóławia się ten podział o stopień odpowiedniości cechy, która dodatkowo może być mocna lub słaba. Ich identyfikacja i weryfikacja są dokonywane za pomocą wybranego algorytmu indukcyjnego. Przyczyną są możliwe skutki heurystycznych technik przeszukiwania. Jak pokazują badania, mogą one prowadzić do sytuacji, w której pojęcie zależy od atrybutów niedecyzyjnych, a w niektórych przypadkach może się nawet zdarzyć, że atrybuty decyzyjne mogą obniżać trafność klasyfikacyjną [20].

Identyfikacja wymienionych rodzajów atrybutów odbywa się na podstawie porównania rozkładów warunkowych poszczególnych wartości atrybutów opisowych i wartości atrybutu pojęcia. Kłopotu z obliczeniem wszystkich warunkowych rozkładów można uniknąć, stosując wybrany algorytm indukcyjny [19].

Wykorzystanie zidentyfikowanych informacji kontekstowych w procesie uczenia się pojęć odbywa się za pomocą jednej lub kilku wybranych strategii. Dostępnych jest pięć strategii, a są to: normalizacja kontekstowa, rozszerzenie kontekstowe, kontekstowa selekcja klasyfikacyjna, kontekstowe dopasowanie klasyfikacyjne, których autorami są A.J. Katz, M.T. Gately i D.R. Collins [21], oraz ważenie kontekstowe autorstwa P.G. Turneya [51; 53].

Celem proponowanych strategii jest eliminacja wpływu kontekstu na cechy kontekstowo zależne, z wyjątkiem ważenia kontekstowego, gdzie mamy do czynienia z preferencją atrybutów mogących poprawić trafność klasyfikacyjną.

Ideą normalizacji kontekstowej zależnej od zadania klasyfikacji jest eliminacja wpływu kontekstu na cechy kontekstowo zależne, którą wykonuje się w fazie przygotowania do procesu uczenia [54; 56; 57]. Polega ona na transformacji atrybutu kontekstowo zależnego na nowy atrybut, którego wartości są wynikiem normalizacji cechy kontekstowo zależnej przez atrybut kontekstowy. Najczęściej stosowanym sposobem jest odjęcie od wartości atrybutu kontekstowo zależnego średniej i podzielenie przez odchylenie standardowe obliczonych dla zakresu wartości wyznaczonego przez dany kontekst. Ograniczeniem tej strategii jest numeryczny typ atrybutów decyzyjnych.

Inną propozycją jest normalizacja kontekstowa niezależna od zadania klasyfikacji, którą przedstawili S. Letourneau i in. [30]. Podejście to, podobnie jak poprzednie, składa się z dwóch kroków: analizy kontekstowej i normalizacji, z tym że atrybuty kontekstowe są znane, a problemem do rozwiązania jest określenie wpływu każdej cechy kontekstowej niezależnie na cechy wykonywalne (atrybuty określające funkcjonowanie danego urządzenia, które są również znane *a priori*). Do analizy kontekstowej zastosowano technikę anova, w której wartości każdego atrybutu kontekstowego dzielone są na przedziały i dokonywana jest zamiana wartości kontekstowych na etykiety przedziałów. Następnie przeprowadza się grupowanie tych zbiorów przedziałów dla każdego atrybutu wykonawczego. Wynikiem analizy kontekstowej jest zbiór tablic opisujących wpływ atrybutu kontekstowego na atrybuty wykonawcze – tablic jest tyle, ile atrybutów kontek-

stowych. Normalizacja następuje poprzez dodanie do wartości atrybutu wykonawczego liczby kontekstowej kary.

Podejście to można szerzej stosować, gdy wykonany zostanie etap identyfikacji cech, chociaż ograniczeniem są rygorystyczne założenia o numeryczności atrybutów wykonawczych.

Niwelowanie wpływu kontekstu na tworzenie klasyfikatora z jednej strony daje możliwość zastosowania nauczonego klasyfikatora do klasyfikacji nowego przykładu występującego w dowolnym kontekście, ale z drugiej strony odrzuca zidentyfikowaną informację kontekstową do dokładniejszego, kontekstowego opisu i interpretacji starych i nowych przykładów.

Kontekstowe rozszerzenie polega na dołączeniu do cech decyzyjnych atrybutów kontekstowych i traktowaniu ich przez klasyfikator w ten sam sposób jak atrybuty decyzyjne. W tym przypadku zadanie nie jest proste, gdyż o decyzyjności atrybutu przesądza heurystyka zastosowanego algorytmu indukcyjnego, chyba że opis pojęcia tworzony jest manualnie.

Kontekstowa selekcja klasyfikacyjna przeprowadzana jest w dwóch krokach. W pierwszym generowane są specjalistyczne klasyfikatory dla każdego zidentyfikowanego kontekstu. W następnym kroku wykorzystuje się cechy kontekstowe do selekcji odpowiedniego klasyfikatora dla określonej sytuacji decyzyjnej. Ograniczeniem jest założenie o jednym kontekście dla klasyfikowanego przypadku.

Kontekstowe dopasowanie klasyfikacyjne jest odwróceniem kolejności kroków w kontekstowej selekcji klasyfikacyjnej. Najpierw przeprowadza się klasyfikację na podstawie cech decyzyjnych, a następnie odbywa się dopasowanie na podstawie cech kontekstowych.

Idea kontekstowego dopasowania jest podobna do dwupoziomowego procesu uczenia zaproponowanego przez R.S. Michalskiego [33] i rozwinięta przez F. Bergadano, S. Matwina i R.S. Michalskiego w systemie POSEIDON [4]. W pierwszej warstwie tworzony jest opis pojęcia niezależny od kontekstów. Jest on generowany z zastosowaniem algorytmu indukcji reguł, a następnie optymalizowany według jakościowych kryteriów opisu, którymi są prostota obliczeniowa i trafność klasyfikacyjna. Proces optymalizacji obejmuje możliwość wprowadzenia zmian w opisie oryginalnym, takich jak usuwanie reguły, warunku lub modyfikacja warunku reguły. Zmiany te odnoszą się mogą zarówno do operacji uogólniania, jak uszczegóławiania oryginalnych reguł.

Druga warstwa ICI (*Inferential Concepts Interpretation*) opisuje możliwe interpretacje pojęcia. Zawiera ona reguły kontekstowe podane przez ekspertów na podstawie analizy błędnie zaklasyfikowanych przypadków z warstwy pierwszej lub nauczone przez klasyfikator na podstawie dodatkowej wiedzy.

Ważenie kontekstowe polega na przypisaniu większej ważności cechom decyzyjnym, które w danym kontekście są bardziej użyteczne w zadaniu klasyfikacji.

Problemem w tym przypadku są wybór atrybutów oraz wysokość przypisanych wag. Dodatkową trudnością może być wielość kontekstów.

Przy założeniu, że kontekst jest zjawiskiem lokalnym w analizowanym problemie, P. Domingos [12] zaproponował technikę selekcji cech kontekstowo zależnych w tzw. leniwym uczeniu (*lazy learning*, *case-based reasoning*, *instance learning*). Jest to sposób podobny do grupowania, polegający na znajdowaniu dla każdego przykładu uczącego przykładu najbardziej podobnego z tej samej klasy i weryfikowaniu hipotezy, że różnice w wartościach atrybutów są nieistotne. Klasyfikacja ma charakter lokalny zależny od konkretnego przykładu w kontekście odpowiedniości klasyfikacyjnej atrybutów.

### 3.2. Uczenie przyrostowe

Zmienność uczonego pojęcia w czasie wymusza zastosowanie innego trybu uczenia, którym jest uczenie przyrostowe. Kontekst w przyrostowym uczeniu się pojęć to zadanie identyfikacji zmian i nieciągłości występujących w procesie uczenia lub między przykładami uczącymi. Zmiany tego typu zwane są dryfem pojęciowym [28; 31]. Zakłada się, że zmiany w środowisku uczenia mogą wpływać na zmianę pojęcia docelowego, którą algorytm uczenia próbuje zidentyfikować w każdym kroku procesu uczenia.

Zmienność pojęcia może mieć charakter stopniowy lub skokowy, różna może też być również prędkość zachodzenia zmian. Nie każde tempo zmian pojęcia jest możliwe do przesledzenia. Z reguły pojawiające się rozwiązania dotyczą powolnych zmian. Określeniem możliwości śledzenia dryfu pojęciowego w teoretycznym modelu uczenia zajmują się m.in. A. Kuh i in. [27], D.P. Helmbold i P.M. Long [16] oraz M. Devaney i A. Ram [11]. Analizują zakres i częstość zmian pojęcia oraz określają szerokość okna niezbędną do efektywnego śledzenia dla stopniowej zmiany pojęcia. Otrzymane wyniki nie mają niestety praktycznego znaczenia w sensie modelu teoretycznego PAC [3; 16].

Rozróżnia się dwa podejścia do radzenia sobie ze zjawiskiem dryfu pojęciowego. Pierwszym jest selekcja przykładów wspomagana przez monitorowanie niezawodności i złożoności hipotez oraz ważenie przykładów, a drugim jest monitorowanie wskaźników sygnalizujących zmiany kontekstów opisywanego pojęcia.

Celem selekcji przykładów jest wybranie adekwatnego zbioru przykładów do bieżącego dryfu pojęcia [61-63], który będzie stanowił podstawę analizy identyfikacji zmian i poprawności hipotez. Do tego celu wykorzystuje się metodę okien, które mogą mieć stały, jak u M. Kubata [24], lub elastyczny rozmiar [42; 63].

Metoda okien polega na uogólnianiu przykładów uczących według zawartości okna przesuwanego wzdłuż przychodzących nowych przypadków. Nauczone pojęcie na podstawie w ten sposób wyselekcjonowanych przykładów wykorzystuje się tylko do przewidywania najbliższej przyszłości.



Elastyczne ustalanie rozmiarów okna polega na monitorowaniu procesu uczenia poprzez bardziej złożone i elastyczne strategie zapominania oparte na różnych heurystykach. Ich celem jest odkrycie zmian w pojęciach docelowych i reakcja na dostrzeżone zmiany poprzez dynamiczną regulację rozmiaru okna. Najczęstszą miarą są czynniki monitorujące trafność predykcyjną [42; 63].

Zamiast selekcji przykładów za pomocą okien M.B. Harries i K. Horn [14] zaproponowali grupowanie kontekstowe. Polega ono na podziale uporządkowanego po czasie zbioru danych na określoną liczbę podzbiorów. Podział może być losowy lub na podstawie wiedzy dziedzinowej. Otrzymane podzbiory są następnie doprecyzowywane według kryterium jakości wygenerowanych opisów (klasyfikatorów). Podejście to ma zastosowanie do sytuacji, w której mamy do czynienia z rozłącznymi i (lub) powtarzającymi się kontekstami.

Innym sposobem determinującym selekcję przykładów jest monitorowanie niezawodności i złożoności hipotez (utworzonych opisów pojęcia). Na przykład w systemie DARLING trafność odnoszona jest do przykładów, które nie zostały wyeliminowane przez bardzo podobny przykład z takim samym lub innym kontekstem pojęcia [41].

Pierwszym z pomysłów monitorowania niezawodności hipotez był algorytm STAGGER Schlimmera i Grangera [43]. Rozważali problem oceny wielu hipotez w odniesieniu do trafności klasyfikacyjnej i wariantów złożoności opisów. Zmianę opisu pojęcia tylko na podstawie śledzenia trafności klasyfikacyjnej klasyfikatora można znaleźć w pracy G. Widmera i M. Kubata [63]. Autorzy, zachowując istniejący opis, analizują możliwość i konieczność tworzenia nowego opisu. Przechowują wcześniejsze modele, które mogą okazać się poprawne w innych sytuacjach.

Inny przykład monitorowania złożoności ewoluujących hipotez można znaleźć w pracach G. Widmera i M. Kubata [58; 61-63], którzy opracowali rodzinę algorytmów FLORA.

Techniką wspierającą selekcję przykładów może być ich ważenie. Odbywa się ono poprzez aktualizację zmian istotności statystycznych w reprezentacji rozkładu opisów pojęcia oraz na podstawie zmniejszającej się wagi przykładów uczących i hipotez wraz z upływem czasu. Dezaktualizacja przykładów definiowana jest przez własność starzenia się, która najczęściej jest funkcją czasu. Z reguły przyjmuje się założenie o większej wiarygodności informacji bieżącej od informacji starszych [22]. Oprócz czasu, ważenie można odnieść do kompetencji opisu bieżącego pojęcia. Wazenie nie zawsze jest w stanie efektywnie wesprzeć selekcję [23]. Liczba analizowanych przykładów w skrajnym przypadku może ograniczać się tylko do przykładów z bieżącego okna, a hipotezy nie wspierające przykładów z tego okna są usuwane [41].

Drift pojęciowy może być rozróżniany poprzez monitorowanie atrybutów wskaźnikowych charakteryzujących środowisko rozpatrywanego pojęcia. Atrybuty wskaźnikowe mogą być znane na podstawie wiedzy dziedzinowej lub zidentyfiko-

wane przez eksperta dziedzinowego. Przykładem pierwszej sytuacji jest m.in. system SAMUEL, którego autorami są C.L. Ramsey i J.J. Grefenstette [40]. Monitoruje on zbiór zdefiniowanych atrybutów wskaźnikowych, który działa w środowisku zmiennym, a jego rozszerzona wersja w środowisku stałym.

Sytuację, gdy atrybuty wskaźnikowe kontekstu nie są znane, rozpatruje G. Widmer [59; 60]. Proponuje on system METAL(B), który automatycznie odkrywa kontekstowe atrybuty wskaźnikowe w procesie uczenia przyrostowego i potrafi odpowiednio reagować na przewidywaną zmianę kontekstu. System składa się z podstawowego poziomu uczenia, który wykonuje regularne uczenie i zadanie klasyfikacji oraz metapoziomu identyfikującego potencjalne wskaźniki kontekstowe (cechy, które mogą być kontekstowymi wskaźnikami). Uczenie kontekstowe i wykrywanie występują w regularnym procesie uczenia bez rozdzielania fazy uczenia i rozpoznawania kontekstu. Wprowadza pojęcia cech oraz atrybutów predykcyjnych i kontekstowych oraz kryteria rozpoznawania kontekstów. Cechy predykcyjne identyfikowane są w odniesieniu do bieżącego okna, ponieważ trafność jest jakością temporalną, która może zmienić się z upływem czasu i kontekstem. O predykcyjności danej cechy decyduje wynik porównania rozkładu wartości danej cechy i bezwarunkowego rozkładu klas w bieżącym oknie. Kryterium to zwraca uwagę na globalne zmiany w rozkładzie między predykcyjnością różnych cech. Natomiast identyfikacja kontekstowości danej cechy dokonywana jest na podstawie tego, jak jej wartości korelują z predykcyjnością innych cech.

Zapewnienie uniwersalności selekcji przykładów uczących, stanowiących podstawę uczenia się pojęcia w wielu różnych kontekstach, wymaga stosowania równocześnie wielu okien o różnych rozmiarach. Problem jest szczególnie ważny w złożonych i heterogenicznych bazach danych. Rozpoznane konteksty mogą stanowić bardziej lub mniej rozpoznane struktury, a więc konieczna jest wiedza o zależnościach między kontekstami. Należy odkryć charakter kontekstów, który może być lokalny, kompatybilny, charakteryzować się hierarchiczną lub płaską strukturą. Bez takiej wiedzy wnioskowanie po kontekstach jest niemożliwe.

#### **4. Reprezentacja pojęcia kontekstowo zależnego**

Celem reprezentacji pojęcia kontekstowo zależnego jest umożliwienie właściwego wyboru rozumienia i interpretacji pojęcia w konkretnej sytuacji. Forma reprezentacji wyznaczana jest przez przedmiot, którym może być łączny opis pojęcia kontekstowo zależnego lub oddzielny opis pojęcia i jego możliwych kontekstów.

Tradycyjne propozycje reprezentacji pojęcia zakładają, że pełne znaczenie kontekstów zawarte jest w pojedynczej strukturze. Przyjmuje się, że taka struktura obejmuje wszystkie powiązane własności pojęcia i określa jego granice [10; 35; 44; 45]. Tego rodzaju reprezentacje wynikają z zastosowanych algorytmów uczenia się pojęć, o których można powiedzieć, że stosują pewną formę selekcji atrybu-

tów kontekstowo zależnych [12]. Należą do nich algorytmy budowy drzew decyzyjnych, indukcji reguł i leniwego uczenia. Każda ścieżka w drzewie decyzyjnym, reguła w zbiorze reguł oraz każda zasada kwalifikacji dla pojedynczego przykładu może być traktowana jako reprezentacja opisu pojęcia w jakimś niezdefiniowanym kontekście. W przypadku drzew i zbiorów uporządkowanych reguł mówi się o zależnych opisach kontekstowych o określonej strukturze.

Niezależne opisy kontekstowe mogą być reprezentowane za pomocą nieuporządkowanego zbioru reguł oraz grafu kontekstowego CxG. Graf kontekstowy jest acykliczny i skierowany, z jednym wejściem i jednym wyjściem. Ścieżka od wejścia do wyjścia reprezentuje wariant kontekstowego opisu pojęcia, czyli wiele różnych opisów w różnych kontekstach. Grupuje wszystkie warianty kontekstowych opisów pojęcia. Jawna reprezentacja opisów umożliwia radzenie sobie z różnymi doświadczeniami zależnymi od kontekstu. Grafy kontekstowe potwierdziły swoją przydatność w wielu dziedzinach, np. w obszarze wspomaganie decyzji w zarządzaniu wypadkami w systemie paryskiego metra [6]. Warunkiem powodzenia jest sterowanie wyborem właściwego opisu pojęcia.

Element kontekstowy w tego typu reprezentacjach może być traktowany jako czynnik aktywizujący określony fragment wiedzy [5]. Jest to sytuacja, w której pojęcie jest widziane w wielu rozłącznych kontekstach. Stąd możliwa jest tylko jedna interpretacja w danym momencie. Ten aspekt kontekstu ma charakter lokalny oznaczający wykorzystanie tylko części wygenerowanego opisu pojęcia, ale bez możliwości nazwania kontekstu.

Dokładniejszy opis i reprezentacja pojęcia kontekstowego są możliwe, gdy konteksty, w jakich występuje pojęcie, są znane i można je oddzielić od opisu podstawowego pojęcia, a konteksty są rozłączne.

Jednym z pomysłów jest dwuwarstwowa reprezentacja pojęcia (opis bezpośredni i pośredni pojęcia) zaproponowana przez R.S. Michalskiego [33]. Pierwsza warstwa BCR (*Base Concept Representation*) opisuje podstawowe charakterystyki pojęcia niezależne od kontekstu w postaci zbioru reguł (pojęcia typowe). Druga warstwa ICI (*Inferential Concepts Interpretation*) opisuje możliwe interpretacje pojęcia. Warstwa ta może mieć postać nie tylko bazy wiedzy, również lecz także innych modeli wiedzy, np. modelu sieci neuronowej [25].

Bardziej uniwersalną reprezentację ze względu na możliwość opisanie dowolnie złożonej struktury kontekstu i pojęcia podstawowego zaproponowali Terziyan i Puuronen [46]. Jest to wielopoziomowa reprezentacja pojęcia kontekstowego za pomocą semantycznej metasieci. Na poziomie pierwszym najniższym znajduje się wiedza o podstawowej dziedzinie pojęcia i jego relacjach. Drugi poziom zawiera relacje metakontekstów, a kolejne poziomy metakonteksty wyższego poziomu itd. Wiedza na najwyższym poziomie jest prawdziwa we wszystkich kontekstach.

Semantyczna metasieć jest zbiorem sieci semantycznych ułożonych jedna nad drugą, tak że relacjami każdego niższego poziomu są węzły poziomu wyższego.

Zadaniem każdej z tych reprezentacji jest wszechstronna obsługa pojęć kontekstowo zależnych poprzez dostarczenie:

- wiedzy zinterpretowanej przez wszystkie znane poziomy jej kontekstu,
- nieznaney wiedzy, gdy interpretacja w jakimś kontekście i sam kontekst są znane,
- nieznaney wiedzy o kontekście, gdy wiadomo, jak wiedza jest interpretowana w tym kontekście.

Do realizacji powyższych zadań zastosowano specjalną algebrę umożliwiającą przeprowadzenie kontekstowych transformacji. Taka reprezentacja jest możliwa, gdy pojęcie występuje w wielu kontekstach charakteryzujących się hierarchiczną organizacją, różną granulacją [7] oraz zagnieżdżeniem [47], a konteksty i ich relacje są znane lub możliwe do rozpoznania.

Oddzielenie kontekstu od podstawowej charakterystyki pojęcia nie zawsze jest możliwe lub uzasadnione, a konteksty nie muszą być rozłączne. Oznacza to, że dane pojęcie może być postrzegane przez pryzmat wielu kontekstów równocześnie. Aspekt ten zwany jest przez Ch. Ghidinię i F. Giunchigilię kompatybilnością [13]. Kompatybilność odnosi się do relacji między kontekstami wspólnego opisu pojęcia. Ma ona strukturalny wpływ na zmianę zbioru lokalnych modeli definiujących każdy kontekst. Kwalifikacja egzemplarza do wygenerowanego opisu pojęcia odbywa się przy określonym współdziałaniu kompetentnych klasyfikatorów. W ekstremalnym przypadku każde dwa konteksty mają dwa niezależne lub takie same spojrzenia na świat.

## 5. Podsumowanie

Przedstawiony przegląd rozwiązań stosowanych w obszarze uczenia się pojęć naturalnych poprzez poszerzenie opisu o możliwe konteksty oraz doświadczenia autorki [18; 19] w tym obszarze badawczym nasuwa kilka spostrzeżeń.

Najpełniejsze możliwości rozpoznania kontekstów daje uczenie wsadowe, gdyż wszystkie konteksty z założenia znajdują się w zbiorze danych. Jest możliwość odkrycia ich struktury i wzajemnych powiązań. Wydaje się, że może ono stanowić podstawę przyrostowego korygowania i uzupełniania kontekstów w procesie uczenia przyrostowego, gdy takie jest konieczne. Nawet w przypadku wymagania uporządkowania danych po czasie, rozpoznanie takie jest możliwe, jak pokazują na przykład badania P. Domingosa [12]. W takim podejściu do uczenia przyrostowego już na wstępie znane byłyby przynajmniej w zarysie możliwe konteksty oraz ich złożoność, co mogłoby ułatwić otrzymanie właściwych opisów pojęcia. Uzasadnieniem jest możliwość rozpoznawania wielu kontekstów w uczeniu przyrostowym, gdy są powiązane relacjami hierarchicznymi, a jest to zadanie trudne i nie podejmowane.

Obszarem wartym rozszerzenia jest również specjalizacja kontekstowa, jako sposób wykorzystania informacji kontekstowej, której poświęcono niewiele uwagi.

Problemem do rozwiązania jest łączenie decyzji klasyfikatorów kontekstowych, gdy konteksty danego pojęcia nie są rozłączne, ale zachodzą na siebie. Selekcja klasyfikatora dla danego przypadku nie jest wówczas właściwym sposobem wyboru opisu. Każdy przypadek byłby opisany przez grupę klasyfikatorów, które wspólnie podejmowałyby decyzję według jakiegoś ustalonego schematu.

W złożonych problemach społeczno-ekonomicznych trudno liczyć na możliwość wyodrębnienia opisu podstawowego i opisu kontekstu. Wydaje się, że dwupoziomowa reprezentacja, choć bardzo atrakcyjna ze względów poznawczych, jest trudna do uzyskania. Nawet jeśli takie rozwiązanie jest możliwe, to nie ma pewności, że otrzymane rozdzielanie jest poprawne. Bardziej uniwersalnym podejściem będzie zwolnienie założenia o generowaniu dwu poziomowego opisu pojęcia. Nie oznacza to jednak, że otrzymane opisy pojęcia w różnych kontekstach nie będą miały wspólnych fragmentów.

Pomijanym problemem jest również przyjmowanie założenia o kompletności opisów pojęcia kontekstowo zależnego (problem szczególnie dotyczy uczenia wsadowego). Przeszkodą dla takiego założenia są bazy danych, stanowiące podstawę uczenia się pojęć. Nie ma gwarancji, że stanowią reprezentacyjną próbkę przykładów analizowanego pojęcia, gdyż były tworzone w innym celu. Nie są wynikiem zaplanowanego eksperymentu. Jeśli tak, to przypadki w nich zawarte nie muszą reprezentować wszystkich zidentyfikowanych kontekstów w takim samym stopniu. Dobrym pomysłem mogłoby być dopuszczenie klasyfikatorów ocenianych nie globalnie, ale według kompetencji dla danej wartości pojęcia. Wówczas zastosowanie miałyby tylko kompetentne i adekwatne fragmenty opisu pojęcia.

## Literatura

- [1] Akman V., *Rethinking context as a social construct*, „Journal of Pragmatics” 2000, Vol. 32, No. 6, s. 743-759.
- [2] Akman V., Surav M., *Steps toward formalizing context*, „AI Magazine” 1996, Vol. 17, No. 3, s. 55-72.
- [3] Barve R.D., Long P.M., *On the complexity of learning from drifting distributions*, „Information and Computation” 1997, Vol. 138, No. 2, s. 101-123.
- [4] Bergadano F., Matwin S., Michalski R. S., Zhang J., *Learning two-tiered description of flexible contexts: The POSEIDON system*, „Machine Learning” 1992, Vol. 8, No. 1.
- [5] Brézillon P., *Some characteristics of context*, „Lecture Notes in Computer Science” 2006, Vol. 4031, s. 146-154.
- [6] Brézillon P., Cavalcanti M., Naveiro R., Pomerol J.-Ch., *SART: An intelligent assistant for subway control*, „Pesquisa Operacional, Brazilian Operation Research Society” 2000, Vol. 20, No. 2, s. 247-268.
- [7] Brézillon P., Gonzalez A., *Tale of two context-based formalisms for representing human knowledge*, „Lecture Notes in Computer Science” 2006, Vol. 4031, s. 137-145.
- [8] Brézillon P., Pomerol J.-Ch., *Contextual knowledge sharing and cooperation in intelligent assistant systems*, „Le Travail Humain” 1999, Vol. 62, No. 3, s. 223-246.

- [9] Caruana R., *Algorithms and Applications for Multi-Task Learning*, Proceedings of the 13<sup>th</sup> International Conference on Machine Learning, Morgan Kaufmann, San Francisco 1996.
- [10] Collins A.M., Quillian M.R., *Experiments on semantic memory and language comprehension*, [w:] L.W. Gregg (red.), *Cognition, Learning and Memory*, John Wiley, New York 1972.
- [11] Devaney M., Ram A., *Dynamically adjusting concepts to accommodate changing contexts*, [w:] *Working Notes*, ICML-96 Workshop on Learning in Context-Sensitive Domains, Bari, Italy, 1996.
- [12] Domingos P., *Context-sensitive feature selection for lazy learners*, „Artificial Intelligence Review” 1997, Vol. 11, No. 1-5, s. 227-253.
- [13] Ghidini Ch., Giunchiglia F., *Local models semantics, or contextual reasoning = locality+compatibility*, „Artificial Intelligence” 2001, Vol. 127, No. 2.
- [14] Harries M.B., Horn K., *Learning stable concepts in a changing world*, [w:] G. Antoniou, A. Ghose, M. Trucszinski (red.), *Lecture Notes on Artificial Intelligence 1359: Learning and Reasoning with Complex Representation*, Springer-Verlag, Berlin, New York 1998.
- [15] Harries, M. B., Sammut, C., Horn, K., *Extracting hidden context*, „Machine Learning” 1998, Vol. 32.
- [16] Helmbold D.P., Long P.M., *Tracking drifting concepts by minimizing disagreements*, „Machine Learning” 1994, Vol. 1, No. 1 s. 17-45.
- [17] Hirst G., *Context as a spurious concept*, [w:] *Paper Presented at the AAAI-97 Fall Symposium on Context in Knowledge Representation and Natural Language*, MIT, Cambridge, Mass., 1997.
- [18] Jakubczyc J., *Contextual classifier ensemble for predicting customer churn*, [w:] M. Nycz, M. Owoc (red.), *Knowledge Acquisition and Management*, Research Papers of Wrocław University of Economics No. 25, Publishing House of the Wrocław University of Economics, Wrocław 2008.
- [19] Jakubczyc J., *Contextual classifier ensembles*, [w:] W. Abramowicz (red.), *LNCS 4439 – Business Information Systems*, Springer, Berlin, Heidelberg 2007, s. 562-569.
- [20] Jon G.H., Kohavi R., Pfleger K., *Irrelevant features and subset selection problem*, [w:] W.W. Cohen, H. Hirsh (red.), *Machine Learning: Proceedings of the Eleventh International Conference*, Morgan Kaufmann Publishers, San Francisco 1994, s. 121-129.
- [21] Katz A.J., Gately M.T., Collins D.R., *Robust classifiers without robust features*, „Neural Computation” 1990, Vol. 2, s. 471-479.
- [22] Kirizakova I., Kubat M., *FAVORIT: Concept formation with ageing of knowledge*, „Pattern Recognition Letters” 1992, Vol. 13, nr 1, s. 19-25.
- [23] Klinkenberg R., *Learning drifting concepts: Example selection vs. example writhing*, „Intelligent Data Analysis” 2004, Special Issue on Incremental Learning Systems Capable of Dealing with Concept Drift, Vol. 8, No. 3, s. 281-300.
- [24] Kubat M., *Floating approximation in time-varying knowledge bases*, „Pattern Recognition Letters” 1989, Vol. 10, s. 223-227.
- [25] Kubat M., *Recycling decision trees in numeric domains*, „Informatica” 2000, Vol. 24, No. 3, Slovenia.
- [26] Kubat M., Holte R.C., Matwin S., *Machine learning for the detection of oil spills in satellite radar images*, „Machine Learning” 1986, Vol. 30, No. 2-3, s. 195-215.
- [27] Kuh A., Petsche T., Rivest R.L., *Learning time-varying concepts*, [w:] *Advances in Neural Information Processing Systems (NIPS)*, Vol. 3, Morgan Kaufmann, San Francisco, Cal., 1991, s. 183-189.
- [28] Langley P., Gennari J.H., Iba W., *Hill-climbing theories of learning*, [w:] *Proceedings of the Fourth International Workshop on Machine Learning*, Morgan Kaufman, Los Altos, Cal., 1987, s. 312-323.
- [29] Lenat D.B., Guha R.V., *Ideas for Applying CYC*, Cycorp, Inc., <http://www.cyc.com/tech-reports/act-cyc-407-91/act-cyc-407-91.html>, 1991 (dostępny 22.12. 2008).
- [30] Letourneau S., Famili A.F., Matwin S., *A normalization method for contextual data: Experience from a large-scale application*, Machine Learning: EMCL-98, „Lecture Notes in Computer Science” 1998, Vol. 1398, s. 49-54.

- [31] Littlestone N., *Learning quickly when relevant attributes abound: New linear-threshold algorithm*, „Machine Learning” 1988, Vol. 2, No. 4, s. 285-318.
- [32] Matwin S., Kubat M., *The role of context in concept learning*, [w:] *Proceedings of the ICML-96 Workshop on Learning in Context-Sensitive Domains*, Bari 1996, s. 1-5.
- [33] Michalski R.S., *Concepts as Flexible and Context-dependent Sets: The Two-tiered View*, Technical Report, Mason Archival Repository Service, 1991, <http://hdl.handle.net/1920/1688> (dostępny 23.03.2009).
- [34] Michalski R.S., *Learning flexible concepts: Fundamental ideas and methods based on two-tiered representation*, [w:] *Machine Learning: An Artificial Intelligence Approach*, Vol. 3, Morgan Kaufmann Publishers, San Mateo, Cal., 1990.
- [35] Minsky M., *A framework for representing knowledge*, [w:] P. Winston (red.), *The Psychology of Computer Vision*, McGraw-Hill, New York 1975.
- [36] Mitchell T., Caruana R., Freitag D., McDermott J., Zabowski D., *Experience with a learning personal assistant*, „Communications of the ACM” 1994, Vol. 37, No. 7. s. 80-91.
- [37] Motschnig-Pitrik R., *An integrated view on the viewing abstraction: Contexts and perspectives in software development, AI, and Databases*, „Journal of Systems Integration” 1995, Vol. 5, No. 1, s. 23-60.
- [38] Palmquist M., *The Role of Context in Shaping Purpose and Constructing Meaning*, <http://writing.colostate.edu/guides/processes/writingsituations/conclusion.cfm> (dostępny 23.03.2009).
- [39] Powell G., Matheus C., Kokar M., Lorenz D., *Understanding the role of context in the interpretation of complex battlespace intelligence*, [w:] *Proceedings of the Ninth International Conference on Information Fusion*, 2006, <http://handle.dtic.mil/100.2/ADA457435>.
- [40] Ramsey C.L., Grefenstette J.J., *Case-based initialization of genetic algorithms*, [w:] *Proceedings of the 5<sup>th</sup> International Conference on Genetic Algorithms*, Morgan Kaufman, San Mateo, Cal., 1993.
- [41] Salganicoff M., *Destiny-adaptive learning and forgetting*, [w:] *Proceedings of the Tenth International Conference on Machine Learning*, Morgan Kaufmann, San Mateo, Cal., 1993, s. 276-283.
- [42] Salganicoff M., *Tolerating concept and sampling shift in lazy learning using prediction error context switching*, „Artificial Intelligence Review” 1997, Vol. 11, s. 133-155.
- [43] Schlimmer J.C., Granger R.H., *Incremental learning from noisy data*, „Machine Learning” 1986, Vol. 1, s. 317-354.
- [44] Smith E.E., Medin D.L., *Categories and Concepts*, Harvard University Press, Cambridge, Mass., 1981.
- [45] Sowa J., *Peircean foundations for a theory of context*, „Lecture Notes in Computer Science”, *Proceedings of the International Conference on Conceptual Structures: Fulfilling Peirce’s Dream*, Springer-Verlag, London 1997, s. 41-54.
- [46] Terziyan V.Y., Puuronen S., *Multilevel context representation using semantic meta network*, „International Journal of Medical Informatics” 1998, Elsevier.
- [47] Theodorakis M., Analyti A., Constantopoulos P., Spyrtos N., *A theory of contexts*, [w:] *Information Bases. Technical Report 216*, Institute of Computer Science Foundation for Research and Technology – Hellas, March 1998.
- [48] Tomaszewski T. (red.), *Pamięć. Uczenie się. Język*, Wydawnictwo Naukowe PWN, Warszawa 1995.
- [49] Tomaszewski T., Grabowska A., Budohoska W., Koziński J., *Percepcja. Myślenie decyzyjne*, Wydawnictwo Naukowe PWN, Warszawa 1995.
- [50] Tsybmal A., Pechenizkiy M., Cunnigham P., Puuronen S., *Dynamic integration of classifiers for tracking concept drift in antibiotic resistance data*, Computer Science Technical Report

- TCD-CS-2005-26, Trinity College Dublin, Department of Computer Science, <https://www.cs.tcd.ie/publications/tech-reports/reports.05/TCD-CS-2005-26.pdf> (dostępny 23.03.2009).
- [51] Turney P.D., *Exploiting context when learning to classify*, [w:] *Proceedings of the Sixth European Conference on Machine Learning*, Springer-Verlag, Berlin 1993.
- [52] Turney, P.D., *The identification of context-sensitive features: A formal definition of context for concept learning*, [w:] *13th International Conference on Machine Learning (ICML96)*, Workshop on Learning in Context-Sensitive Domains, Bari, Italy, July 1996, s. 53-59.
- [53] Turney, P.D., *The management of context-sensitive features: A review of strategies*, [w:] *13th International Conference on Machine Learning (ICML96)*, Workshop on Learning in Context-Sensitive Domains, Bari, Italy, July 1996, s. 60-66.
- [54] Turney P.D., Halasz M., *Contextual normalization applied to aircraft gas turbine engine diagnosis*, „Journal of Applied Intelligence” 1993, Vol. 3, No. 2, s. 109-129.
- [55] Wang Y., Joshi M., Rose C., Fisher F., Weinberger A., Stegmann K., *Context based classification for automatic collaborative learning process analysis*, [w:] *Frontiers in Artificial Intelligence and Applications. Proceeding of the Conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work*, 2007, Vol. 158, s. 662-664.
- [56] Watrous R.L., *Speaker normalization and adaptation using second-order connectionist networks*, „IEEE Transaction on Neural Networks” 1993, Vol. 4, No. 1, s. 21-30.
- [57] Watrous R.L, Towell G., *A patient-adaptive neural network ECG patient monitoring algorithm*, *Computers in Cardiology*, Siemens Corp. Res. Inc., Princeton, NJ, 1995. s. 229-232.
- [58] Widmer G., *Combining robustness and flexibility in learning drifting concepts*, [w:] *Proceedings of the Eleventh European Conference on Artificial Intelligence*, Wiley, Chichester 1994.
- [59] Widmer G., *Recognition and exploitation of contextual clues via incremental meta-learning*, [w:] *Proceedings of the 13<sup>th</sup> International Conference on Machine Learning*, Morgan Kaufmann, San Francisco, Cal., 1996.
- [60] Widmer G., *Tracking context changes through meta-learning*, „Machine Learning Journal” 1997, Vol. 27, No. 3, s. 2755-2790.
- [61] Widmer G., Kubat M., *Effective learning in dynamic environments by explicit context tracking*, [w:] *Proceedings of the Sixth European Conference on Machine Learning*, Springer-Verlag, Berlin 1993.
- [62] Widmer G., Kubat M., *Learning flexible concepts from streams of examples: FLORA2*, [w:] *Proceedings of the 10<sup>th</sup> European Conference on Artificial Intelligence*, Wiley, Chichester, UK, 1992.
- [63] Widmer G., Kubat M., *Learning in the presence of concept drift and hidden contexts*, „Machine Learning” 1986, Vol. 23, No. 1, s. 69-101.

## CONTEXT IN CONCEPT LEARNING

**Summary:** Growing volumes and complexity, big number of sources and goals of gathering data that are different than data analysis tasks are the main reason of failure of concept learning algorithms. Simplistic and inadequate to the reality assumptions become significant. They concern accuracy of concepts, contextual independency and possibility to represent concept by a single nominal description. This paper presents an overview of ways to unchain those assumptions and broadening description of concepts by contextual dependency.