

SPRAWOZDANIE Z DZIAŁALNOŚCI NAUKOWO-BADAWCZEJ KATEDRY INFORMATYKI STOSOWANEJ UMK W ROKU 2007

I. TEMATY BADAWCZE

1. Liczba realizowanych tematów.

- 14, w tym 9 w ramach działalności statutowej, 5 w ramach badań własnych.

2. Liczba tematów zakończonych: 0 - w planowanym terminie 0, - z opóźnieniem 0.

- Prowadzone badania mają w większości charakter długoterminowy.

3. Z liczby tematów zakończonych przekazano:

- do wykorzystania, - do wdrożenia 0.

4. Z liczby tematów przekazanych do wdrożenia, w 2007 r. wdrożono:

- ogółem 6, - w tym w resorcie edukacji narodowej 0.

- Nie wdrożono wyników nowych tematów ale dokonano pewnych ulepszeń w pakiecie programów GhostMiner 3 do data mining, opracowanym przy współpracy z firmą FQS Poland (oddział Fujitsu) i wdrożonym w poprzednich latach; obecnie powstaje całkiem nowa wersja systemu do analizy danych w oparciu o prowadzone badania.

5. Liczba tematów, których realizację zaniechano (w roku sprawozdawczym).

- nie zaniechano żadnego tematu.

6. Liczba realizowanych grantów KBN

2 granty własne:

- Grant KBN „Selekcja informacji i odkrywanie struktur w danych”, lata 2005-2007, kierownik W. Duch.
- Grant KBN „Meta-uczenie w inteligencji obliczeniowej”, lata 2005-2007, kier. N. Jankowski.
- Granty marszałka woj. Kujawsko-Pomorskiego dla dwóch doktorantów (2007).
- Udział w Grancie KBN „Biblioteka InfoSel++ (Information Selection Library in C++) jako narzędzie wstępnego przetwarzania danych” kierowanym przez J. Biesiadę z Politechniki Śląskiej.

II. CZĘŚĆ OPISOWA

II.1. Przyczyny zaniechania tematów prowadzonych i rozpoczętych w 2007 roku.

W 2007 roku nie zaniechano realizacji żadnego tematu.

II.2. Krótka charakterystyka i podsumowanie wyników uzyskanych w 2007 r.

A) Inteligencja obliczeniowa (computational intelligence)

Główne osiągnięcia w tej grupie tematów to:

1. W Springer "Studies in Computational Intelligence" Series, 2007 ukazała się książka: W. Duch i J. Mandziuk, Challenges for Computational Intelligence. Książka składa się z kilkunastu rozdziałów na temat wyzwań stojących przed inteligencją obliczeniową, napisanych przez najlepszych specjalistów w tej dziedzinie. Liczący 55 stron rozdział książki „Towards comprehensive foundations of computational intelligence” podsumowuje nasze prace nad meta-uczeniem.
 - Duch W, Mandziuk J (Eds.), Challenges for Computational Intelligence. Springer "Studies in Computational Intelligence" Series, Vol. 63, June 2007, 488 pp.
2. 9-13 września odbyła się w Porto 17th International Conference on Artificial Neural Networks (ICANN 2007), oficjalna doroczna konferencja European Neural Network Society. W. Duch jest od 2005 roku prezydentem tej organizacji, na konferencji w Porto został wybrany na kolejną 3-letnią kadencję. Pełnił również rolę „Program Co-chair” (z prof. D. Mandicem z Imperial College, Londyn) i był współredaktorem materiałów w objętości ponad 2000 stron wydanych w dwóch tomach w ramach Springer Lecture Notes in Computer Science, Vol. 4668 i 4669.
 - Marques de Sá, J, Alexandre L.A, Duch W, and Mandic D.P. (Eds.), Proceedings of the 17th International Conference on Artificial Neural Networks (ICANN 2007), Part I, Springer Lecture Notes in Computer Science, Vol. 4668, 2007, XXXI, 978 pp. ISBN 978-3-540-74689-8.
 - Marques de Sá, J, Alexandre L.A, Duch W, and Mandic D.P. (Eds.), Proceedings of the 17th International Conference on Artificial Neural Networks (ICANN 2007), Part II, Springer Lecture Notes in Computer Science, Vol. 4669, 2007, XXXI, 990 pp. ISBN 978-3-540-74693-5.
3. W 2007 r. realizowano grant KBN na rozwój metod selekcji informacji. Dr J. Biesiada rozwijał pakiet programów realizujący znane i nowe metody przydatne do dyskretyzacji i selekcji cech. Na temat selekcji informacji napisano szereg prac prezentowanych na wielu międzynarodowych konferencjach. Kilka nowych pomysłów dotyczących stabilizacji jakości wybieranych cech jest w realizacji.
 - Biesiada J, Duch W, Feature Selection for High-Dimensional Data: A Pearson Redundancy Based Filter. Advances in Soft Computing Vol. 45, 242-249, 2008.
4. W 2007 r. realizowano też grant KBN na rozwój metod meta-uczenia, polegających na automatycznym wyborze najlepszych alternatywnych modeli dla danego punktu złożoności/dokładności. Prowadzono prace nad nowym pakietem Intemi, który stanowić będzie nową generację programów data mining, zdolnych do automatycznego odkrywania interesujących modeli danych. Ukazały się pierwsze prace na temat jego konstrukcji.
 - Grąbczewski K. i N. Jankowski, Versatile and Efficient Meta-Learning Architecture: Knowledge Representation and Management in Computational Intelligence, IEEE Symposium Series on Computational Intelligence, IEEE press, April 2007, pp. 51-58
 - Duch W, Towards comprehensive foundations of computational intelligence. In: W. Duch and J. Mandziuk, Challenges for Computational Intelligence. Springer Studies in Computational Intelligence, Vol. 63, 261-316, 2007.
5. Rozwijano dalej idee k-separowalności, która pozwala na charakteryzację problemów nieseparowalnych liniowo na klasy o różnej złożoności. Wprowadzono nowe cele uczenia i nowe algorytmy, które powinny pozwolić na uczenie się w przypadkach trudnych, z którymi nie radzi sobie obecnie żaden istniejący algorytm. Odkryto w ten sposób bardzo proste i dokładne modele. Obecnie reali-

zowanych jest kilka wersji algorytmów dla k-separowalności. Dokonano też porównań tego podejścia z mało jeszcze znanymi metodami uczenia geometrycznego, osiągając wysoce konkurencyjne rezultaty.

- Grochowski M, Duch W, Learning highly non-separable Boolean functions using Constructive Feedforward Neural Network, Lecture Notes in Computer Science, Vol. 4668, 2007, pp. 180-189
6. Rozwijano nadal nowe metody logicznego opisu danych oparte na szukaniu prototypów i optymalizacji funkcji odległości, w tym probabilistycznych miarach podobieństwa, konkurencyjne do podejścia opartego na systemach nerorozmytych. Wysłano rozdział do książki o używaniu algorytmów SVM do szukania reguł logicznych. Mgr Blachnik ukończył i obronił z wyróżnieniem pracę doktorską oraz pracował przez 4 miesiące w laboratorium prof. Erkki Oja w Helsinkach.
 - Blachnik M, Duch W, Prototype rules from SVM. Rozdział do książki, ed. J. Diederich, Springer 2007
 7. Rozwijano badania związane z analizą tekstów medycznych i ogólnymi metodami analizy tekstów; opracowano szkic neurokognitywnego podejścia do rozumienia tekstów, integrujące pamięć rozpoznawczą, semantyczną i epizodyczną, zaczęto tworzenie sieci semantycznych z konkurencją zawierających wiedzę a priori pozwalającą na interpretację tekstu, grafów spójnych koncepcji, oraz opracowano nowy algorytm stosujący programowanie liniowe do klasyfikacji tekstów. Kilka prac już opublikowano, kolejne są w przygotowaniu. Przeprowadzono międzynarodowy konkurs w tej dziedzinie, zakończony na konferencji IEEE CIS w kwietniu 2007.
 - Duch W, Matykiewicz P, Pestian J, Neurolinguistic Approach to Vector Representation of Medical Concepts. Proc. of the 20th Int. Joint Conference on Neural Networks (IJCNN), Orlando, IEEE Press, August 12-17, 2007, pp. 3110-3115
 - Itert L, Duch W, Pestian J, Influence of a priori Knowledge on Medical Document Categorization, IEEE Symposium on Computational Intelligence in Data Mining, IEEE Press, April 2007, pp. 163-170.
 8. W ramach współpracy z 'Laboratory for Neuroinformatics' (RIKEN - Brain Science Institute (BSI), Japonia), prowadzono badania z grupą prof. S. Usui na następujące tematy:
 - Zastosowano algorytm Spherical Embedding (SE) do wizualizacji w 3 wymiarach danych tekstowych związanych z Neuroinformatyka. Została zwizualizowana 1) struktura laboratoriów w RIKEN BSI Team map, 2) Visiome Platform: baza artykułów, danych, programów i innym dokumentów związanych z badaniami nad systemem wzrokowym, 3) abstrakty posterów przedstawionych w 2006 roku na największej konferencji w dziedzinie neuronauk: Society for Neuroscience Annual Meeting.
 - Badano właściwości algorytmu SE, tj. określono jak parametry algorytmu oraz cechy bezpośrednio zależne od danych wpływają na jakość rozmieszczenia danych w przestrzeni reprezentacyjnej (tj. punktów na sferach w 3 wymiarach).
 - Prowadzono klasteryzacje dokumentów w celu automatycznego wykrywania lub ekstrakcji naukowych dziedzin i tematów (topics) omawianych w tych dokumentach na podstawie samych ich abstraktów i tytułów.

B) Metody komputerowe w fizyce, chemii i naukach o życiu.

Prof. J. Wasilewski i dr S. Zelek prowadzili badania w ramach tematu: „Struktura energetyczna małych cząsteczek”, zajmując się wybranymi aspektami teorii funkcjonałów gęstości (DFT – Density Functional Theory) i jej zastosowaniami do obliczeń struktury energetycznej i widm elektronowo-oscyłacyjnych cząsteczek, uczestnicząc też w badaniach kierowanych przez prof. K. Jankowskiego z Instytutu Fizyki UMK, dotyczących tej tematyki. Ponadto dr S. Zelek współpracował z Zespołem Teoretycznej Biofizyki Molekularnej kierowanym przez prof. W. Nowaka z Instytutu Fizyki UMK, w ramach tematu komputerowej analizy mechanizmu reakcji hydrolizy nityli. W szczególności:

1. Kontynuowano badania relacji między przestrzeniami orbitalnymi otrzymywanymi w metodach: Hartree-Focka, Bruecknera, CASSCF i w różnych wariantach teorii funkcjonałów gęstości (DFT). Przedmiotem szczególnego zainteresowania była ewolucja indeksów podobieństwa przestrzeni orbitalnych wzdłuż krzywych potencjalnych cząsteczek dwuatomowych, na przykładzie izoelektronowej

serii CO, NO+, CN-. Pokazano, że w dość szerokim otoczeniu geometrii równowagowych przestrzenie orbitali Kohna-Shama zachowują bliskie podobieństwo do przestrzeni orbitali Bruecknera i mają istotny charakter wielokofiguracyjny. Dzięki temu w ramach różnych wariantów DFT uzyskuje się znacznie bardziej realistyczny niż w metodzie Hartree-Focka opis energetyki oscylacji molekularnych. Podobieństwo przestrzeni orbitalnych zachowuje się w przypadku anionów, natomiast jest zmniejszone dla kationów, w wyniku kontrakcji orbitali spowodowanej nadmiarowym ładunkiem dodatnim jąder.

- J. Wasilewski, S. Zelek, Evolution of orbital spaces along potential curves for diatomic molecules. A comparative study of Hartree-Fock, Kohn-Sham, Brueckner and multiconfigurational orbital spaces, publikacja w przygotowaniu.

2. Kontynuowano badania związane z poznaniem mechanizmu reakcji hydratacji nityli, w których przy pomocy obliczeń kwantowomechanicznych, próbowano określić który z zaproponowanych przez Hanga i innych (Structure 5, 691 (1997)) mechanizmów reakcji hydrolizy jest najbardziej prawdopodobny. Korzystając z sugestii recenzentów artykułu obliczono współczynniki PED w poszczególnych drganiach badanych cząsteczek

- L. Peplowski, K. Kubiak, S. Zelek, W. Nowak, A comparative DFT study of substrates and products of industrial enzyme nitrile hydratase, International Journal of Quantum Chemistry – po wprowadzeniu poprawek praca jest na etapie drukowania.

Wyniki wspomnianych badań mechanizmu reakcji hydratazy nityli (w zastosowaniu do metylonitrylu) zostały umieszczone również w pracy magisterskiej:

- Adrian Jasiński, Komputerowe modelowanie hydratacji nityli, Toruń 2007 (opiekun S. Zelek)

3. Podjęto badania nad dynamiką ruchu protonu i deuteronu w układzie mostków wodorowych w klastrach wody i ciężkiej wody, z wykorzystaniem metod funkcjonałów gęstości i uwzględnieniem anharmoniczności. Wstępne wyniki weszły do pracy magisterskiej:

- G. Markowski, Badanie klastrów wody i ciężkiej wody metodami mechaniki kwantowej, Toruń 2007 (opiekun J. Wasilewski)

Dr hab. J. Meller i dr R. Adamczak prowadzili badania nad zastosowaniem metod bioinformatyki i biologii obliczeniowej do problemu rozpoznawania struktur i funkcji białek, metod dopasowania sekwencji w biologii molekularnej, analizy podobieństwa genomów (ang. synteny) a także identyfikacji markerów stanów chorobowych i fenotypicznych w danych klinicznych i genomicznych dotyczących polimorfizmów i profilów ekspresji genów. W szczególności:

1. Dokonano oceny przydatności metod przewidywania obszarów transmembranowych do przewidywania białek GCPR.

- R. Adamczak, Ł. Peplowski and W. Nowak, Performance of Neural Networks Based Transmembrane Helix Prediction Methods Applied to Mosquito Anopheles Gambiae Gprotein Coupled Odorant Receptors, in Applications of Statistical and Machine Learning in Bioinformatics; Advances in Computational and Systems Biology, eds. J. Meller and W. Nowak, Peter Lang GmbH, 2007..

2. Rozpoczęto badania nad wpływem redukcji wymiarowości macierzy PSSM na przewidywania struktur drugorzędowych oraz solwatacji.

3. Kontynuowane są badania nad uogólnioną metodą przewidywania genów, która pozwoli zastosować model wytrenowany na jednym z organizmów grzybów do innych podobnych organizmów.

4. Przeprowadzono badania profili ekspresji genów w chorobach reumatologicznych dzieci. Wyniki zostały opisane w następującej pracy:

- J. Meller, R. Adamczak, M. P. Scola, M. Barnes, S. D. Thompson, M. H. Passo, H. I. Brunner, D. N. Glass, and A. A. Grom; Machine Learning Analysis of Expression Profiles of Synovial Tissue Cytokines Helps Identify Patients with Systemic Onset Juvenile Rheumatoid Arthritis, in Applications of Statistical and Machine Learning in Bioinformatics; Advances in Computational and Systems Biology, eds. J. Meller and W. Nowak, Peter Lang, w druku.

5. Opracowano nową metodę badania oddziaływań międzybiałkowych:

- Porollo A, Meller J. Prediction-based fingerprints of protein-protein interactions. Proteins, 2006 Dec 6; [Epub, na okładce wydania pierwszego numeru w 2007 roku]

6. Przeprowadzono szereg zaawansowanych symulacji procesów biologicznych, ściśle przy tym współpracując z doświadczalnikami.
7. Ukazała się obszerna praca: Duch W, Swaminathan K, Meller J, Computational Intelligence Approaches to Rational Drug Design and Discovery. *Current Pharmaceutical Design*, Vol. 13(14), 1497-1508, 2007.

C) Modelowanie kognitywne (cognitive modeling).

W 2007 roku W. Duch był członkiem grupy EU COST Action B27 "Electric neuronal oscillations and cognition" (ENOC), biorąc w 2007 roku udział w spotkaniu "Neuroscience today" we Florencji, został też członkiem grupy EU COST Action "Advanced Methods for the Estimation Of Human Brain Activity and Connectivity" biorąc w 2007 roku udział w spotkaniu i warsztatach "Neuromath" w Rzymie. Rezultatem jest bardzo ambitny projekt globalnej analizy EEG lub ogólnie wielowymiarowych układów dynamicznych.

Wspólnie z prof. J.G. Taylorem pracował nad wydaniem książki „Roadmap to human level intelligence”, zawierającej najbardziej zaawansowane modele typu architektury kognitywnej. Napisano artykuł zawierający przegląd istniejących architektur kognitywnej na pierwszej konferencji „Artificial General Intelligence”, która odbędzie się w Memphis w marcu 2008; dłuższy rozdział do książki jest w przygotowaniu. Trwają prace nad całkiem nową, wzorowaną na mózgu architekturą kognitywną.

Napisano kilka pionierskich prac na temat intuicji, wglądu, wyobraźni i kreatywności. Będą to tematy, które w najbliższych latach się staną się bardzo ważne.

- Duch W, Intuition, Insight, Imagination and Creativity. *IEEE Computational Intelligence Magazine* 2(3), 40-52, 2007.
- Duch W, Creativity and the Brain. *A Handbook of Creativity for Teachers*. Ed. Ai-Girl Tan, Singapore: World Scientific Publishing 2007, pp. 507-530.
- Duch W, Pilichowski M, Experiments with computational creativity. *Neural Information Processing – Letters and Reviews*, Vol. 11(4-6), 123-133, 2007.

Neurokognitywne podejście do symulacji procesów tworzenia nowych słów zastosowano w praktyce w publicznie dostępnym serwerze. Przeanalizowano zachodzące w mózgu procesy w czasie rozumienia słów, w szczególności eksperymenty psychologiczne z wykorzystaniem torowania, w których widoczne są efekty rezonansu stochastycznego. Jest to pierwszy krok w stronę autentycznych możliwości twórczych komputerów. Rezultaty przedstawiono na konferencjach, w rozdziałach książkowych i wysłanej do druku publikacji.

- Duch W, Matykiewicz P, Pestian J, Towards Understanding of Natural Language: Neurocognitive Inspirations. *Lecture Notes in Computer Science* 4669, 953–962, 2007.

Rozwijano ideę wektorów opisu koncepcji (Concept Description Vectors) w zastosowaniu do reprezentacji semantyki i tworzenia pamięci semantycznej. Rozwijano algorytmy zbierania informacji do pamięci semantycznych z encyklopedii, słowników i tekstów bez struktury, dodając mechanizmy aktywnego uczenia oraz inicjowania aktywnych dialogów w celu zdobycia brakującej informacji. Program do gry w 20 pytań zintegrowany z mówiącą głową i rozpoznawaniem oraz syntezą mowy pracuje obecnie w wersji zarówno lokalnego programu jak i interfejsu do aplikacji sieciowych. Wyniki zostały opisane w kilku pracach.

- Szymanski J, Sarnatowicz T, Duch W, Towards Avatars with Artificial Minds: Role of Semantic Memory. *Journal of Ubiquitous Computing and Intelligence*, American Scientific Publishers.

Wspólnie z Katedrą Neurologii Dorosłych AM w Gdańsku, Nat-Syn-Drugs(Włochy), University of Sheffield (W. Brytania), Politechniką Gdańską, University of Copenhagen, Brain Science Institute, Riken, Japonia, oraz firmami INNOVA i Explora S.r (Włochy) złożono we wrześniu 2007 roku projekt "Quantitative System Investigation Of Acute Stroke Events" (QUASAR), w ramach Collaborative Projects, medium-scale focused research projects 7 Programu Ramowego EU. W części, za którą jesteśmy odpowiedzialni planujemy opracowanie parametrycznego modelu generatora rytmów oddechowych w pniu mózgu i analizę wpływu jego uszkodzeń na rytmy oddechowe, która powoli na modelowanie danych spirograficznych. Ukazała się praca na temat modeli komputerowych problemów neurologicznych i demencji.

- Duch W, Computational Models of Dementia and Neurological Problems. Chapter 17 of a book: Neuroinformatics, C.J. Crasto (Ed), "Methods in Molecular Biology" series (J. Walker, series ed.), Humana Press, Totowa, NJ, pp. 307-336, 2007

III. PUBLIKACJE

III.1. W 2007 roku ukazało się ogółem 48 prac, w tym:

Liczba prac opublikowanych w 2007	48
w tym:	
a) Publikacje z listy filadelfijskiego IIN	13
b) nie dotyczy	
c) Publikacje w innych recenzowanych pismach	1
e) Redakcja książek naukowych, tłumaczenia	4
Rozdziały w książkach	10
Artykuły, recenzje, popularno-naukowe i dydaktyczne	20

III.2. Liczba prac przyjętych do druku: 10

III.3. Przykłady publikacji związanych z działalnością naukową KIS:

Najważniejsze prace (max 5).

1. Duch W, Matykiewicz P, Pestian J, Towards Understanding of Natural Language: Neurocognitive Inspirations. Lecture Notes in Computer Science 4669, 953–962, 2007.
Grant KBN (658-F)
2. Duch W, Pilichowski M, Experiments with computational creativity. Neural Information Processing – Letters and Reviews, Vol. 11(4-6), 123-133, 2007.
Grant KBN (658-F)
3. Duch W, Swaminathan K, Meller J, Computational Intelligence Approaches to Rational Drug Design and Discovery. Current Pharmaceutical Design, 13(14), 1497-1508, 2007.
Grant KBN (658-F, WD), badania statutowe i grant NIH (JM)
4. Porollo A, Meller J; Prediction-based Fingerprints of Protein-protein Interactions, Proteins: Structure, Function and Bioinformatics, 66(3), 630-645 (2007)
Badania statutowe i grant NIH
5. Porollo A, Meller J, Versatile annotation and publication quality visualization of protein complexes using POLYVIEW-3D. BMC Bioinformatics 8:316 (2007)
Badania statutowe i grant NIH

Pełny spis znajduje się w Załączniku 2b.