**Cluster 0:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| Neural networks (NN) have become established as powerful tools for complex pattern recognition problems. One application which appears well suited to NN methods is the identification of prognostic groups, to be used for treatment planning. For many cancer, studies of cancer cell biology have added many factors of potential prognostic value, but the way in which these interact with known factors is generally not well studied. The potential of NNs to model these data in a non-linear fashion has only begun to be explored. NNs are not part of standard statistical packages, making them relatively inaccessible to many statisticians. More importantly, current NN methods cannot accommodate censored outcome variables.  This proposal is for development of algorithms for censored-data NNs, implementation of these within a comprehensive statistical package, and evaluation of alternative approaches. The aim is to provide statisticians involved with clinical decision making with more ready access to NN technology, and with the means to analyze survival-type data. The value of NNs in this field cannot be addressed by any single investigator, but by providing the software that is needed, and some guidelines for its use, we anticipate that research in this field will be stimulated. | https://reporter.nih.gov/project-details/2272927 | R41 | 1994 | 0.9 |
| The goal of this research is to develop improved techniques, both fully  automated and computer-assisted, for classification of medical text. The  technical approach is exemplar-based: robust information retrieval methods  find similar, previously-classified texts, and corresponding codes are  used to suggest likely classifications for a new text. Phase I focused  upon implementing experimental software to establish baseline performance  with several variations of the exemplar-based approach.    Phase II builds upon this work to implement a complete Coder's Workstation  (CWS). Based upon Phase I results and assessments of commercial  opportunities, Phase II will focus upon shorter texts (<12 words), which  are best suited for automated methods. A "short-similarity" capability  will be added to the Phase I approach to further enhance performance with  shorter texts. To evaluate and refine the CWS, Phase II will include  extensive "beta testing" of the software at the Brigham and Women's  Hospital and the Mayo Clinic.    The major technical innovation of this project is the development of  highly automated classification software that is sensitive to term  similarities. The major health-related contributions are large potential  savings in coding expenses, reduced time demands upon physicians for  coding, and improved consistency in classification of free text for  research studies.    PROPOSED COMMERCIAL APPLICATION: The proposed technology will have  important commercial application within hospitals, insurance companies,  and pharmaceutical companies which currently expend significant resources  on coding of free text (ICD9, CPT4, COSTART, etc.). The founders of  Belmont Research Inc. have extensive experience in creating and marketing  software to support biomedical applications. | https://reporter.nih.gov/project-details/2429835 | R44 | 1997 | 0.9 |
| In Phase I of the "Decision Support System to Identify the At-Risk  Fetus" a prototype Intelligent Decision Support System (IDSS) was  developed to interpret electronic fetal monitoring (EFM) data. In  Phase II, the researchers propose to complete development and test  the IDSS in preparation for full clinical trials and  commercialization. The research is based on the hypothesis that  intrapartum EFM and the resulting FHR and UC data provide  information that can assist physicians in more accurately  differentiating between healthy and at-risk fetuses. The system  uses morphological filters to process signals at different scales,  a neural network to better recognize FHR and UC patterns despite  EFM noise, and a fuzzy relational structure outcome inferencing. In  Phase I evidence was generated supporting the hypothesis when the  IDSS with little training on only 50 cases differentiated between  healthy and at-risk fetuses as well, and perhaps slightly better  than three board certified obstetricians with over 40 years of  combined clinical experience. The proposed Phase II improvements to  IDSS would increase the system's sensitivity and specificity so as  to assist clinicians in more accurately identifying at-risk  fetuses, as well as more confidently identifying healthy fetuses so  as to avoid unnecessary interventions.    PROPOSED COMMERCIAL APPLICATIONS The successful completion of Phase  II should lead to a system which can be used in hospitals and  clinicians' offices to better identify the at-risk fetus, thereby  allowing for the timely implementation of efficacious interventions  and a reduction in the use of invasive obstetrical interventions,  including cesarean deliveries, without adversely impacting  outcomes. The use of this system in hospitals and private  clinicians' offices will also help reduce the costs of national  health care. | https://reporter.nih.gov/project-details/2421880 | R44 | 1997 | 0.9 |
| DESCRIPTION:(Adapted from applicant's abstract): Implementation and  validation of a computerized expert system sleep scoring is proposed.  Such a system would increase productivity of sleep laboratories, and  increase the quality while reducing the cost of sleep diagnostic  procedures. Phase I research will focus on the basic algorithms for  automated sleep stage scoring. The algorithms are based on Bayesian  decision theory and follow the steps of the human expert in the decision  making process when scoring sleep by the Rechtschaffen-Kales standard.  Likelihood estimates of the suggested and alternative decisions will be  used to quantify confidence in suggested scores. The system will flag  low confidence scores to be revised by human interaction. Initial tests  show that the level of agreement between machine and human is similar  to that between two human scorers. Software tools will be developed to  test the effect of the parameters of the algorithms on stage score  decisions. An initial data base of training and test sets of normal and  apneic sleep recordings will be established and the algorithms will be  optimized and tested on this data. Phase II will optimize parameters and  validate the system on larger data sets for other diagnostic and age  groups, develop on-line, real-time capabilities, and extend the decision  model to facilitate further research in sleep medicine. | https://reporter.nih.gov/project-details/2703049 | R44 | 1998 | 0.84 |
| DESCRIPTION: Access to comprehensive medical information is literally  a matter of life and death for health professionals. The applicants  propose to develop a state-of-the-art medical information system based  on natural language processing (NLP) which provides innovative access  first to the literature of complementary medicine and then to the  traditional medical literature. The goal of Phase I is dual: to  develop a finely tuned, immediately usable information system for  alternative medicine, and to determine on an experimental level what  linguistic elements within the medical subject domain are critical to  optimizing retrieval. For this purpose, an extensive analysis of  medical sublanguage and the text structure of medical documents will be  undertaken, and the results will be applied to each module of the system  in order to optimize it for medical domain. The Phase I system will be  used by the medical community both for access to hard-to-find  information, and to begin to assess the usefulness of complementary  medicine techniques in treating chronic problems. Phase II will add  traditional medical literature to the system to provide a fully  integrated solution for rich, precise access to medical information.    PROPOSED COMMERCIAL APPLICATION: NOT AVAILABLE | https://reporter.nih.gov/project-details/2867891 | R43 | 1998 | 0.84 |

**Cluster 1:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
|  | https://reporter.nih.gov/project-details/2237591 | F37 | 1995 | 1 |
| The long term goal of our research is to understand the flow of  information from the genome to the phenotype of organisms. In this  proposal, we will attempt to use Bayesian networks and near-optimal  sequence alignments to represent protein secondary structures and motifs.  A Bayesian network describes the likelihood of amino acids at each  position in a motif as well as the dependence of amino acids in one  position on the amino acids at other position. Hence, Bayesian networks  can describe both the conservation of amino acids at single positions and  the conservation of correlations between two positions simultaneously.    Conserved amino acids result from evolutionary selection for a specific  amino acid or type of amino acid at one position in a protein structure.  These positions often have important functional or structural  requirements. Correlated changes between amino acids generally result from  side-chain side-chain interactions between pairs of amino acids in a  protein's structure. The types of correlations we have represented with  Bayesian networks include electrostatic charges, hydrophobicity, hydrogen-  bond donor and acceptor and inversely correlated packing volumes among  others. These Bayesian networks can be used to 1) discover side-chain  side--chain interactions within protei motifs and 2) to search sequence  databases for motifs showing both correlations and conserved amino acids.    Near-optimal alignments between two sequences can display regions that  have been more highly conserved or less highly conserved using the  information contained in only two sequences. The most highly conserved  region correspond to the most highly structured regions and the most  highly variable regions correspond to loops and coils and other  hypervariable regions. We propose to use near-optimal alignments to  display conserved secondary structures of proteins and hypervariable  regions. We will use secondary-structure specific amino acid substitution  matrices to provide specificity.    The goals of this proposal are to 1) build a database of Bayesian networks  that represent protein motifs, 2) test these networks for their ability to  detect motifs using test sets and crossvalidation methods, 3) compare  these networks with other methods for searching protein databases , 4)  build an integrated set of Bayesian networks to predict protein secondary  structure, 5) compare the prediction of protein secondary structure with  existing method 6) build a near-optimal sequence alignment workbench, and  7) predict structured and unstructured regions in proteins from near-  optimal alignments. | https://reporter.nih.gov/project-details/2238097 | R01 | 1995 | 1 |
| Abstractions of time-stamped clinical data are useful for planning  therapy, for monitoring therapy, and for creating high-level summaries of  time-oriented clinical databases. Temporal abstractions also support  explanations by an intelligent patient-record system and can be used for  representation of the goals and intentions of clinical guidelines and  protocols.    We propose to reengineer and expand the scope of the RESUME system, a  prototype computer program that implements the knowledge-based temporal-  abstraction method, a conceptual and computational framework that we have  developed for abstraction of time-stamped clinical data into clinically  meaningful interval-based concepts. RESUME has been evaluated with highly  encouraging results in several clinical areas. We will address the  practical and theoretical issues of representation, acquisition,  maintenance, and reuse of temporal-abstraction knowledge. Our specific  aims are defined by a four-step research plan:    1. We will define formally the knowledge requirements for five  computational modules (mechanisms) we employ, thus facilitating the  acquisition, maintenance, reuse, and sharing of the required knowledge.    2. We will enhance, expand, and redesign five computational temporal-  abstraction mechanisms:  (a) Automatic formation of meaningful contexts for interpretation of  clinical data.  (b) Classification of clinical data that have equivalent time stamps into  higher-level concepts.  (c) Temporal inference (e.g., the join of certain interval-based clinical  abstractions into longer ones).  (d) Interpolation between temporally disjoint clinical abstractions,  including a development of a probabilistic representation and semantics.  (e) Matching of predefined and runtime temporal patterns, given time-  stamped data and conclusions.    3. We will develop a tool for automated acquisition, from expert  physicians, of temporal-abstraction knowledge, using techniques from the  PROTEGE-II project for designing knowledge-based systems.    4. We will validate and evaluate our methodology and its implementation.  (a) We will assess the value of the knowledge-acquisition tool in several  experiments.  (b) We will validate the performance of the computational mechanisms in  the domain of therapy of patients who have insulin-dependent diabetes by  collaboration with expert endocrinologists.  (c) We will evaluate the overall framework within EON, a project in which  researchers are implementing an integrated architecture for protocol-based  care. | https://reporter.nih.gov/project-details/2415719 | R29 | 1997 | 1 |
| This Small Business Innovation Research Phase I project will develop  several software products dealing with the analysis of stochastic point  processes for use in basic electrophysiology and behavioral neuroscience,  as well as in fields dealing with higher aspects of human behavior such as  mental health and alcoholism treatment relapse issues. Although not  specifically targeted, because of the ubiquity of the subject matter, we  anticipate considerable usage by researchers in Epidemiology and other  medical fields, and even in diverse fields such as electrical engineering,  forensics and operations research. The software will focus on the  analysis of single and paired event data, i.e., single or multiple lists  of the times that events of a particular type occur. The algorithms  selected for inclusion in the program will include some of the better-  known existing algorithms for point processes, as well as selected,  recently developed algorithms in this area. The software will be for the  PC platform and distributed in the form of a stand-alone windows  application, as well as DLL's and scripts to allow inclusion of the  functionality in existing software products.    PROPOSED COMMERCIAL APPLICATION: The proposed software will be a research  tool that should have wide marketability within the target research  communities. In addition, because of the complete lack of comparable  software this software has the potential to penetrate the much larger non-  specialist market. | https://reporter.nih.gov/project-details/2675191 | R44 | 1998 | 1 |
| The development of a prototype software package, the Gen-Pep Algorithm,  an entirely new, receptor sequence, hydrophobic free energy  eigenfunction-based proprietary system for short peptide design, will  be completed. Its feasibility with respect to a yield of new peptides  which elicit target receptor-mediated biological activation in receptor  cDNA transfected cell lines will be tested using Cytosensor  Microphysiometry. Gen-Pep exploits a sequence of linear  transformations of the hydrophobic free energies of the amino acid  sequences of peptide and their receptors and includes eigenvalue  decomposition and complex pole spectral and wavelet transformation and  a constructive step using the leading eigenvectors. Given a peptide  receptor sequence, this system generates a family of candidate peptides  to interact with it. As exemplars, short peptide analogues of  neurotensin and cholecystokinin will be designed and tested for their  recently demonstrated direct action on the D2 dopamine receptor and the  dopamine transporter proteins respectively. New short peptide analogues  of (longer peptide) epidemoid growth factor, transforming growth factor  and fibrobast growth factor will be designed and tested similarly.  Licensing and support of Gen-Pep including a superfamily hydrophobic  free energy eigenfunction library, contract new peptide development for  biotechnology, companies and new peptide-drug development constitute the  three commercial goals of the enterprise.    PROPOSED COMMERCIAL APPLICATION  1) Licensing Gen-Pep software, support and eigenfunction library for an  annual fee to biotechnology and pharmaceutical research concerns. 2)  Contracting with similar entities for Cielo Institute, Inc. to generate  and preliminarily test promising candidate, short peptides designed for  specific purposes. 3) Developing and testing new peptide ligands for  patenting by Cielo Institute, Inc. | https://reporter.nih.gov/project-details/6021666 | R43 | 1999 | 1 |

**Cluster 2:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| Neural network optimization algorithms greatly enhance our ability to  construct large-scale, dynamical models of highly interconnected networks.  Until now, optimization has only been applied to networks of simplistic  processing units, ignoring the integrative and temporal response  properties of single neurons, thus limiting the predictive power of the  models. The long-term goal of this project is to develop a hybrid  modeling strategy in which optimization methods are applied to networks of  realistic,multicompartmental model neurons. To accomplish this goal, we  will construct a hybrid model of an actual distributed processing network  composed of repeatably identifiable sensory, motor, and interneurons that  computes a well-defined behavioral input-output function. Optimization  will be used to predict the connectivity of as-yet-unidentified  interneurons in the actual network and the predictions will be tested by  identifying the interneurons by physiological and morphological means.  Performance of the hybrid model will be assessed by comparing it to the  performance of an a priori model in which all connection strengths are  determined physiologically. The final model will be used to predict the  loci of synaptic plasticity underlying nonassociative conditioning of the  reflex by incorporating local learning rules and by optimization methods.  The predictions will be tested by determining the actual plastic sites  physiologically. This project will have the combined effect of enhancing  the predictive power of optimized network models and illuminating the  relation between computations at the single-neuron and network levels. | https://reporter.nih.gov/project-details/2250635 | R29 | 1995 | 1 |
| This proposal describes the development and evaluation of a rule-based  inductive machine-learning system to aid in the development of clinical  decision-support aids from data. This system will be specifically  designed to allow a clinical expert to add domain knowledge to improve  the learning process.    Physicians often use rules as clinical guidelines in their everyday  practice. Because individual rules are so easily understood, we believe  that a rule-based system is ideally suited to the enhancement of domain  knowledge for the learning of clinical decision-support systems.    Many clinical databases are sufficiently small and/or biased as to  prevent current statistical and machine-learning methods from producing  optimal clinical decision-support aids. We believe this system will  improve the learning of rule-based models for clinical decision-support,  especially from these inadequate databases.    A secondary aim of this project is to improve on a current rule-based,  inductive machine-learning system by adding global ruleset evaluation  metrics to the rule learning process. | https://reporter.nih.gov/project-details/2237607 | F37 | 1996 | 1 |
| The public's perception of the risk of HIV infection by blood transfusion  has greatly intensified the concern regarding safety of the blood supply,  thus, making transfusion safety a matter of national priority. A major  constraint in transfusion safety is that there is no means for the  systematic collection and analysis of indicents of transfusion medicine  errors. However, incident reporting systems have been developed in other  error-critical fields including aviation, nuclear power, and  anesthesiology. These systems can be used as guiding models for the  development of an similar reporting system in transfusion medicine.  Therefore, the specific aims of this project are to: (l) design a  prototype reporting system for the collection and classification of  incidents with the potential for compromising the safety of the blood  supply, (2) develop and construct an operational prototype reporting  system based upon the design criteria, (3) demonstrate the effectiveness  of the reporting system in collecting, storing, and classifying  information related to safety and human error at multiple sites through  implementation testing, (4) derive rational strategies for enhancement of  human performance and safety based upon the analysis of the classification  of error types contained in the prototype data system, (5) evaluate the  prototype reporting system's effectiveness, document the development  process, and report project outcomes. This project will be carried out as  an interdisciplinary effort involving experts from the fields of  transfusion medicine, education and training, cognitive psychology,  artificial intelligence, aviation safety, and nuclear power. This will be  achieved using consensus development. The system will be implemented and  tested in three blood centers (Blood Care of Dallas, Dallas, Texas; New  York Blood Center, New York City, NY; Oklahoma Blood Institute, Oklahoma  City, OK) and three hospital transfusion services (Parkland Memorial  Hospital, Dallas, TX; New York University Medical Center, New York City,  NY; University of Southern California Medical Center, Los Angeles, CA).  This prototype system may well serve as a national model for improving  safety of the nation's blood supply. | https://reporter.nih.gov/project-details/2029311 | R01 | 1997 | 1 |
| DESCRIPTION (Adapted from applicant's abstract): The "closed-loop  artificial pancreas," a device that would measure glucose level and  deliver insulin automatically as needed, has been an elusive goal in the  treatment of diabetes. There are three essential components: the blood  glucose sensor, linking algorithms and the delivery system. For the first  time, a viable sensor and a proven delivery system are now available for  research. The broad goal of this clinical research proposal is to complete  the studies needed to link the sensor to the delivery system, paving the  way for a functional closed-loop artificial pancreas. First, we will make  a detailed analysis of sensor signal as it reflects glucose level in  normal and diabetic humans. Second, we will study the precise  pharmacokinetics of insulin delivery by external and implantable insulin  pumps. Third, analysis of these two data sets will provide the basis for  algorithms that link the sensor signal to insulin delivery. A formal  safety analysis will evaluate the safety features needed in a closed loop  device. In the last year of the project, the entire system will be tested  and fine-tuned. This project takes advantage of our relatively extensive  investigational experience with mechanical insulin delivery pumps in  people with diabetes, and the recent availability, for research, of a  subcutaneously placed, glucose oxidase-based continuous glucose sensor.  The investigators have established experienced with clinical research in  diabetes, and the resources of an excellent General Clinical Research  Center. The co-investigators have extensive experience with mathematical  modeling of biologic systems. There is a close working relationship  between the research team and the manufacturer of the sensor and pumps, as  reflected by the Interactive Research Project Grant collaboration, and by  a long-standing history of collaboration. It is essential to emphasize  that we do not anticipate completion of a manufacturable, clinically  usable, commercially viable artificial pancreas within the time-frame of  this work. Rather, we aim to complete the basic studies and modeling  analyses that would form the basis of such a system, and demonstrate the  feasibility of linking the sensor to the delivery device. If these studies  and these trials were successful, they would be a major step towards  development of a clinically useful close-loop artificial pancreas. | https://reporter.nih.gov/project-details/2759047 | R01 | 1998 | 1 |
| This project aims to develop and commercialize significantly improved software for digital enhancement of the detail of chromosome banding patterns in microscopic images. These investigators have developed an innovative technique for this application, based upon wavelet transforms and multiresolution image analysis. Used with modern computerized chromosome analysis the proposed technique promises significantly improved enhancement of chromosome banding patterns and more effective visual detection of subtle rearrangements. This will help clinicians and researchers detect previously invisible or sub-visible band pattern alterations in conventional and high resolution banding. It will significantly increase the ability of automated instruments to assist the evaluation of chromosome alterations in clinical samples and in normal and neoplastic mammalian cells. During Phase I we implemented and tested three wavelet transforms with desirable mathematical properties. We developed a prototype multiresolution image processing system for chromosome enhancement. We obtained extremely encouraging results, strongly suggesting that these techniques offer considerably improved enhancement capability over conventional methods. and clearly demonstrating the feasibility of this approach. In Phase II we will complete the implementation and refinement of the software. We will implement several wavelet design approaches and evaluate many wavelet transform basis function sets that potentially can bring out relevant detail in chromosome banding patterns. PROPOSED COMMERCIAL APPLICATIONS: As soon as the new enhancement techniques are developed and qualified for routine application, they will be incorporated into PSII's PowerGene products, both in new systems sold and as an upgrade to existing systems. | https://reporter.nih.gov/project-details/2793651 | R44 | 1999 | 1 |

**Cluster 3:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The electronic medical record (EMR) holds great allure to both the  medical informatics and health services research communities. In this  project, we propose to enhance the capability of electronic medical  record (EMR) systems by creating and evaluating tools to extract  clinical vocabularies as well as patient data from narrative text  reports. We will apply advanced natural language processing tools from  the CLARIT system to both of the above problems. We contend that fast  and robust automated text processing methods are the only way that the  problems of vocabulary construction and narrative text extraction can  be solved.    We will address the clinical vocabulary problem by utilizing the  thesaurus extraction techniques already present in the CLARIT system.  Using several gigabytes of narrative text, including discharge  summaries, progress notes, radiology reports, and other clinical text,  we plan to:  l. Identify empirically the terminology used in medicine.  2. Compare the coverage of that terminology in several existing large  medical vocabularies: UMLS, SNOMED, and the Medical Entities  Dictionary.  3. Discern the semantic characteristics of that terminology to allow  other structured vocabularies a richer substrate of terms as well as  providing us the opportunity to implement a clinical vocabulary schema  based on the methods of the MedSORT-II Project.  4. Evaluate how well our tools assist the vocabulary building efforts  of ourselves and others.    The narrative extraction problem will be approached differently than  in the past, building on the efforts of previous investigators who  have tackled this problem before but changing the perspective by  focusing on the development of tools specific to researchers and  others with a need to extract data from narrative text. This approach  will be applied in two domains:  l.Consortium-based research in the use of esophogastroduodenoscopy  (EGD).  2.Practice guidelines implementation in blood product transfusion. | https://reporter.nih.gov/project-details/2332614 | U01 | 1997 | 0.87 |
| The research described in this proposal begins with completion and formal  evaluation of MIDAS "a computer program designed to automatically  construct decision models from an underlying medical knowledge base". The  capabilities of MIDAS will be extended to knowledge-based construction  of Markov decision models. A second project will develop a comprehensive  knowledge management scheme for the problem of pulmonary disease in AIDS.  This scheme will use a knowledge base structured according to knowledge  needed to perform a decision analysis. It will incorporate summaries of  relevant data, sources and quality of data and links to the original  sources. This knowledge management scheme will be deployed in the  hospital and evaluated in a group of medical residents at Robert Wood  Johnson University Hospital. | https://reporter.nih.gov/project-details/2546277 | K04 | 1997 | 0.87 |
| LONG-TERM OBJECTIVES 1. Develop a computerized system, based on  hierarchical neural network pattern recognition technology, for reliable  identification of plants. 2. Identify poisonous plants. 3. Expedite  discovery of new medicinal plants. 4. Create an image database directly  from plant material and link with existing medicinal plant databases.  5. Develop commercial product for pharmaceutical companies, agriculture  and others. SPECIFIC AIMS 1. Design hierarchical system of neural  networks to follow natural plant taxonomy groupings and extend our  identification technology to a large number of plant species. 2.  Improve accuracy of identification. 3. Design a prototype workstation  for botanical and agricultural field stations and laboratories.    RESEARCH DESIGN AND METHODS FOR ACHIEVING GOALS. 1. Digitize large  number of plant species from special collections. 2. Measure  automatically venation patterns and shape. 3. Design hierarchical  neural networks to divide plants into natural groupings. 4. Accumulate  virtual herbarium database as leaves are digitized (scanned or  photographed).    POTENTIAL FOR TECHNOLOGICAL INNOVATION This system is unique in  capturing botanical recognition knowledge in a hierarchy of neural  networks and is the first fully-computerized system for plant  identification utilizing information digitized directly from plants.    PROPOSED COMMERCIAL APPLICATION  1. Expedite discovery of new medicinal plants for pharmaceutical  industry. 2. Create valuable database directly from plants. 3.  Identification of poisonous plants. 4. Valuable for rapid  identification of invasive weeds. | https://reporter.nih.gov/project-details/2644669 | R44 | 1998 | 0.87 |
| DESCRIPTION: (Applicant's abstract) The feasibility of a Drowsiness  Monitoring Device (DMD) to detect EEG indices of drowsiness in real-  time, was demonstrated during Phase I with an analytical model correctly  classifying 97.1 percent of sleep episodes and 94.3 percent of awake  epochs in 20 sleep-deprived subjects. The model employs discriminant  function analysis (DFA) to characterize and classify one-sec epochs,  validated against a combination of visual scoring by polysomnographers  and/or a behavioral measure. Algorithms to detect artifacts in real-time  (EMG, 60-Hz and gross body/eye movements) were developed. This  classification accuracy represents a significant advancement over  previously reported models and confirms the feasibility of  distinguishing EEG characteristics of sleep and waking on a second-by-  second basis. Phase II will implement a three-level DFA classification  system to further refine the model, adding sub-class states for  vigilance and drowsiness/sleep to improve system accuracy. The multi-  dimensional time-series DFA analyses will correlate EEG parameters with  behavioral measures of driving performance to provide quantitative  predictions of performance decrements associate with sleep onset.  The model will be validated using a population with demographics  consistent with the target market for the DMD (e.g., truck drivers)  during Phase II. In addition, the system will be evaluated with the  introduction of commonly used legal drugs (caffeine, nicotine, and cold  medications) to determine the robustness of the model.    PROPOSED COMMERCIAL APPLICATION:  The DMD provides three levels of user safety. When sleep onset is  approaching, the DMD will initiate a verbal warning alarm that must be  turned off by the user. Alternatively, the DMD can provide verbal  feedback to ensure the user maintains high levels of alertness during  activities that require sustained vigilance. The user can also select  the option for the DMD to recommend the optimal time to take a short  nap, monitor the length of the nap and awaken the user at the  appropriate time.    Currently, more that 10% of the U.S. workforce or an estimated 20  million people are engaged in night sift work. The transportation  industry, including airline, railroad, marine and highway transportation  companies, is the nation's third largest employer of shift workers. Long  haul truck drivers, in particular, are vulnerable to sleepiness because  they drive through the night, in most cases unaccompanied, and generally  sleep less than 6 hours per day at irregular intervals. In addition, an  estimated 6 million Americans suffer from chronic sleep disorders which  make them vulnerable to fatigue in the workplace. | https://reporter.nih.gov/project-details/2794305 | R44 | 1999 | 0.87 |
| The overall goal of this continuing project is to develop efficient  algorithms which will permit the construction of realistic and  comprehensive mathematical models that relate normal and pathological  renal function to the underlying membrane transport and flow processes in  the renal tubules and their associated vasculature. The primary thrust of  our research during the next period will be:    1. To use our present inner medullary models to develop fairly detailed  architectural models of the inner stripe of the outer medulla and then  integrate these new models with our present central core and vasa recta n-  nephron inner medullary models.    2. To include additional solutes and osmolytes and investigate the role of  osmolyte production in the inner medulla.    3. To incorporate a more detailed representation of transmural movement of  water and solutes that takes into account both cellular and paracellular  pathways and also the exchange of electrolytes and water between red blood  cells and plasma in the vasa recta.    4. Adapt current algorithms for serial computers, - if necessary, develop  new ones based on our split system solvers - to fully exploit the parallel  and vector processing capabilities of supercomputers. This is necessary to  handle the size and complexity of our current and future n-nephron models.    5. If necessary - for parallel and/or vector algorithms - (a) improve  stability and accuracy of numerical methods, (b) develop hierarchal  solution strategies, and (c) incorporate continuation and smoothing  methods.    6. To make the non-linear Schur Complement type methods developed by us  readily available to other biomedical modelers, for use on minicomputers  and/or workstations. These include: (a) reduced models of the whole kidney  and medulla, (b) models of epithelia and isolated perfused tubules, (c)  tubuloglomerular feedback response models, and (d) neural network models. | https://reporter.nih.gov/project-details/2137100 | R01 | 1995 | 0.48 |

**Cluster 4:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The purpose of this contract is to provide research and development support for the Unified Medical Language System. Specifically, the contractor will support (1) the definition of the functional components and algorithms needed: to relate the user's terms to information in the UMLS Knowledge Sources; to interact with the user to clarify the information needed; and to select and access the information sources relevant to the user's inquiry, and (2) the evaluation of the utility of the UMLS Knowledge Sources and proposed functional components in a variety of environments. | https://reporter.nih.gov/project-details/2319081 | N01 | 1992 | 1 |
| The primary objective of this Data Analysis grant proposal is to employ  and evaluate the ability of Neural networks to serve as clinical tools for  predicting clinical coronary events. Several methods will be employed to  adapt neural networks for survival analysis, and data from town ling-term  clinical studies will be used to develop neural networks which evaluate  coronary artery disease risk. The Specific Aims for this proposal are:    Specific Aim 1 (Relation between Angiographic Progression of Coronary  Artery Atherosclerosis using Quantitative Coronary Angiography and  Clinical Coronary Events in CLAS): Neural networks will be developed  utilizing quantitative coronary angiography (QCA) measures of coronary  change.    Specific Aim 2 (Relation between Risk Factors and Serum Lipid Values and  Clinical Coronary Events in the LRCCPPT): Neural networks will also be  developed utilizing information of serum Lipid values available in the  Lipid Research Clinics Coronary Primary Prevention Trial results (LRCCPPT  Results, 1984a and Appendix B).    Specific Aim 3 (Evaluation of Methods which Adapt neural Networks to  Survival Analysis): Existing neural Network designs cannot readily  incorporate the censored observations which exist in long-term clinical  studies. Methods developed by our research team (Lapuerta, Azen and  LaBree, 1995; and Buckley and James, 1979) provide approaches which can be  used to address this problem. | https://reporter.nih.gov/project-details/2029573 | R03 | 1996 | 1 |
| The overall objective of this Phase I SBIR effort is to design a  personally customized, interactive multimedia expert system for  promotion of exercise programs among chronically ill elderly patients,  called the "Personalized Interactive Exercise" ("PIE"). This will be  built as a specialized module on top of a comprehensive,multimedia  distributed network computer system for "Continuously Available Medical  Care" (CAMC) focused on the needs of chronically ill patients. Using the  advanced CAMC Tool Kit scripting and expert system language. physicians  will input medical condition parameters and patients will provide  personal preferences and home/living factors, then the PIE software will  create a customized training and monitoring CD-R0M which accounts for  medical problem factors such as diabetes and peripheral vascular  disease. Multimedia video training, rules for progression of exercise  and live videoconferencing to physicians and exercise trainers will be  features of the PIE system. As part of this Phase I effort, a prototype  will he constructed and tested at congregate housing for the elderly in  conjunction with researchers from the Hebrew Rehabilitation Center for  the Aged in Boston.    PROPOSED COMMERCIAL APPLICATION: The market for PIE modules and products  is substantial. The marginal value of PIE to the CAMC homecare system  is approximately $100 per year; for the 1 million high cost illness  patients this totals $100 million per year as a market. Even if LG&A  reaches only several percent, this is very attractive annual recurring  revenue base with high net income after the first few years. Sales of  PIE "authoring" software for physicians to create PIE CD-ROMs for their  patients, priced at $750 (as an example) per primary care physician to  a market with 300,000 potential buyers is also a several hundred million  dollar opportunity; again a small percent of this market is an  attractive business for LG&A. | https://reporter.nih.gov/project-details/2002237 | R43 | 1997 | 1 |
| DESCRIPTION (Adapted from applicant's abstract): This is a relatively  old problem, that so far has been essentially intractable. The  applicants propose to use a "threading technique" augmented by further  use of some other methods together with neural networks or weighted  procedures. All procedures proposed have previously been used, and the  applicants will try to use several of them with threading as the basic  technique, and a weighted voting scheme to help select what hopefully  will be the correct structure.    PROPOSED COMMERCIAL APPLICATION: NOT AVAILABLE | https://reporter.nih.gov/project-details/2332123 | R43 | 1998 | 1 |
| DESCRIPTION (adapted from investigator's abstract): The aim of this Phase  II project is to test the Resident-Centered Information (RCI) system for  Assisted Living: a service planning format using an expert decision  system, an integrated resident-services logging system, and a complete set  of management modules. The system supports assisting living providers  serving the frail elders and other dependent groups and includes  approaches to meet the needs of those with cognitive impairment. This  computerized system is easy to use so that it is supportive to staff and  creates reliable data for continuous improvement and outcome research.    PROPOSED COMMERCIAL APPLICATION: NOT AVAILABLE | https://reporter.nih.gov/project-details/6055449 | R44 | 1999 | 1 |

**Cluster 5:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| There is growing recognition that, while practice guidelines can  predispose physicians to behavioral change, even the most respected and  clinically relevant guidelines often do not affect actual practice.  Despite the apparent reluctance of physicians to incorporate published  practice guidelines into their clinical routine, there are signs that  immediate automated reminders can have a profound influence on physician  behavior, and on the cost and quality of medical care. As practice  guidelines proliferate and become more complex, the dilemma facing those  who design practice guidelines is the tradeoff between the  comprehensiveness of the guidelines and the likelihood they will be  retained and implemented by the physicians for whom they are intended.  This same complexity can complicate computer-based implementation of  guidelines. For example, the dependence of guidelines on patient  preferences and individual risk factors makes the rule-based approach  impractical for many guidelines.    Intelligent decision systems (IDS) are specifically designed to implement  complex, flexible guidelines tailored to varying clinical circumstances.  These systems provide advice based on a tailored decision-analytic model.  But these systems assume that a physician-user will be persuaded by a  review of the decision model and its associated quantitative results.    The goal of this work is to develop and test improved explanation methods  for quantitative decision models so that intelligent decision systems can  be used in a consultative mode. We have developed a program called QxQ,  which uses symbolic reasoning to provide qualitative text explanations for  the results of decision trees. This work is intended to extract the key  architecture- and domain-independent elements of QxQ's methodology, and to  reimplement and test them as part of MIDAS, an existing IDS. First, we  will develop a domain- and system-independent version of QxQ, called EQxQ  (for essential QxQ), by applying it to decision models constructed in  MIDAS. In collaboration with the developers of MIDAS, we will identify the  key independent elements necessary to generate explanations from an IDS.  Second, we will devise and implement additional explanation methods in  EQxQ to allow explanation of complex modeling constructs, such as Markov  models, cycle trees, and cost-effectiveness models. Third, we will test  the robustness of the MIDAS-EQxQ system by using it to implement a  decision model underlying a clinically relevant practice guideline.  Finally, we will establish and pilot test the routine use of the program  in an outpatient clinic. We will survey the physician-users to determine  the strengths and weaknesses of the program. Using pilot data from the  program's routine use, we will design a prospective clinical evaluation of  the system and its explanations by measuring their effect on patient  outcome. Our study will provide data regarding the clinical utility of  intelligent decision systems. | https://reporter.nih.gov/project-details/2430874 | R29 | 1997 | 0.71 |
| Neural networks (NN) have become established as powerful tools for complex  pattern recognition problems. One application which appears well suited to  NN methods is the identification of prognostic groups, to be used for  treatment planning. For many cancer, studies of cancer cell biology have  added many factors of potential prognostic value, but the way in which  these interact with known factors is generally not well studied. The  potential of NNs to model these data in a non-linear fashion has only begun  to be explored. NNs are not part of standard statistical packages, making  them relatively inaccessible to many statisticians. More importantly,  current NN methods cannot accommodate censored outcome variables.    This proposal is for development of algorithms for censored-data NNs,  implementation of these within a comprehensive statistical package, and  evaluation of alternative approaches. The aim is to provide statisticians  involved with clinical decision making with more ready access to NN  technology, and with the means to analyze survival-type data. The value of  NNs in this field cannot be addressed by any single investigator, but by  providing the software that is needed, and some guidelines for its use, we  anticipate that research in this field will be stimulated.    PROPOSED COMMERCIAL APPLICATION: The power of NNs has been recognized in  the marketplace, and NNs are widely used. However, there is a real need to  make NN more accessible for clinical applications by incorporating an easy-  to-use NN program, providing the most commonly needed NN models and  functions, into a mainstream statistical package. Furthermore, none of the  currently available packages addresses the specific problem of censored  data; we expect to find an immediate market among statisticians dealing  with clinical data. | https://reporter.nih.gov/project-details/2460603 | R42 | 1997 | 0.71 |
| The proposed research aims to: a) Improve the understanding of the  genetics of inherited diseases with unclear modes of transmission.  Studies will evaluate the effectiveness of current methods of analysis,  including classical linkage analysis, sib-pair or affected-pedigree-  member analysis and the use of measures of association in understanding  the underlying genetic mechanisms of such traits. Simulation studies  will continue to provide a source of family data reflecting confounding  factors thought to be a problem in linkage analysis of certain complex  traits, such as psychiatric or behavioral disorders. Factors to be  considered include assortative mating, genetic heterogeneity and multi-  locus disease determination. The ability of current methods to  correctly analyze traits with one or more of these factors will be  assessed and, where appropriate, alternative methods will be developed  and tested. b) Apply techniques of neural network pattern matching to  problems of genetic systems. Applications include: aid in phenotype  definition for traits with multiple clinical problems of genetic  systems. Applications include: aid in phenotype definition for traits  with multiple clinical characteristics; determination of risk of disease  based on phenotype, known risk factors and disease profiles in  relatives; determination of organ transplant success based on HLA  antigen matching patterns; definition of disease phenotype based on  quantitative factors. c) Develop and apply strategies for ordering  multiple linked loci using pairwise recombination data, radiation hybrid  data, or other physical mapping data. Some of these ordering strategies  may be adaptable to the development of techniques for integrating map  information obtained by different methods, an important step in  organizing a comprehensive, reliable map. d) Carry out classical  linkage analysis for specific genetic diseases. Currently, a genome  scan is underway to identify a gene or genes for polycystic liver  disease. Other diseases to be studied include lymphoma and prostate  cancer. Methods to be tested in the simulation studies can be applied  to these analyses in order to better understand the complete genetic  picture, including identification of heterogeneity, by detecting linkage  of different disease forms to different marker loci. Such  differentiation will help sharpen the clinical definition of various  forms of the diseases.    As a result of advances from this work, better mathematical tools for  the study of diseases with complex or ill-defined inheritance patterns  will be available. Applications to specific diseases will increase  understanding of interactions between clinical definition and  predisposing genetic factors. This will increase the precision of  genetic counseling and suggest useful approaches for studying the  mechanisms involved in determining disease state. | https://reporter.nih.gov/project-details/2693221 | R01 | 1998 | 0.71 |
| We propose a low-cost commercial intelligent hypermedia information system to educate patients and the general public about breast cancer. User centered system design, focusing on seeing the domain from the patient's or potential patient's point of view, is central to our approach.  The need for breast cancer education evident. About 182 thousand women will be diagnosed with breast cancer this year in the United States alone. Patients-and the general public-want information. Doctors and nurses are pressed for time, and for any number of reasons may not be able to provide all of the information an individual may want or need.  One technological innovation in our proposal is the application of user- centered system design to a commercially available medical information system. A second is the use of cognitive models of patients' understanding and perceived information needs regarding breast cancer and its treatments to develop the deep structure of the interface between the medical model and the user. | https://reporter.nih.gov/project-details/2104339 | R43 | 1994 | 0.57 |
| The goal of this research is to develop improved assessment protocols  that afford a quantitative and analytic evaluation of speech impairment  in children and adults with neurological disorders. Speech impairments  (dysarthrias) will be studied in persons with amyotrophic lateral  sclerosis, Parkinson's disease, stroke, cerebellar degeneration, cerebral  palsy, and developmental speech disorders. Improved evaluation of speech  intelligibility is a particular focus of this work, but issues of speech  and voice quality also are addressed. The methods to be used are a  combination of standard clinical assessments (such as rating scales),  intelligibility evaluations, and computer-based acoustic analyses.  Specifically, the methods include: perceptual ratings of speech by  experienced clinicians, quantitative assessment of intelligibility, a  multiple-parameter acoustic analysis, computer correction of speech  abnormalities through LPC resynthesis of the acoustic signal, and  derivation of vocal tract shape from acoustic parameters. Work in all  these areas will be based on recordings of speech samples from a large  number of individuals with dysarthria. One product of the research will  be a library of clinical profiles including intelligibility scores,  phonetic feature analyses, ratings of speech/voice quality, acoustic  measures, and neurological diagnosis. Particular attention will be given  to the influences of subject age and sex on the characteristics of  dysarthria for a given neurological diagnosis. The assessment protocols  will be implemented on microcomputers and designed to be incorporated in  clinical practice. The research also will contribute to the development  of expert systems for the rating and classification of dysarthria. | https://reporter.nih.gov/project-details/2443579 | R01 | 1997 | 0.57 |

**Cluster 6:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The long-term objective of our research group is to facilitate automatic  or semi-automatic classification and retrieval of natural language texts,  in support of reducing the cost and improving the quality of computerized  medical information. This proposal develops further and applies a novel  approach, the Linear Least Squares Fit (LLSF) mapping, to document  indexing and document retrieval of the MEDLINE database. LLSF mapping is  a statistical method developed by the PI for learning human knowledge  about matching queries, documents, and canonical concepts. The goal is to  improve the quality (recall and precision) of automatic document indexing  and retrieval, which cannot be achieved by surface-based matching without  using human knowledge or thesaurus-based matching dependent on manually  developed synonyms. This project applies LLSF to MEDLINE, the world's  largest and most frequently used on-line database, to evaluate the  effectiveness of this method and to explore the practical potential on  large scale databases. The specific aims and methods are:    l. To collect data needed for the training and evaluation of the LLSF  method. A collaboration with another research institute is planned for  utilizing and refining a large collection of MEDLINE retrieval data. A  sampling of MEDLINE searches at the Mayo Clinic will be employed for  obtaining additional tasks.    2. To develop automatic noise reduction techniques for improving both the  accuracy of the LLSF mapping and the efficiency of the computation. A  multi-step noise reduction in the training process of LLSF will be  investigated, including a statistical term weighting for the removal of  non-informative terms, a truncated singular value decomposition (SVD) for  reducing the noise at the semantic structure level, and the truncation of  insignificant elements in the LLSF solution matrix for noise-reduction at  the level of term-to-concept mapping.    3. To scale-up the training capacity for enabling the LLSF to accommodate  the large size of MEDLINE data. A split-merge approach decomposes a large  training sample into tractable subsets, computes an LLSF mapping function  for each subset, and then merges the lcal mapping functions into a global  one.    4. To improve the computational efficiency by employing algorithms  optimized for sparse matrices and for noise reduction. The potential  solutions include the Block Lanczos truncated SVD algorithm which can  reduce the cubic time complexity of standard SVD (on dense matrices) to a  quadratic complexity, a QR decomposition which solves the LLSF without  SVD, a sparse matrix algorithm which has shown a speed-up in matrix  multiplication and cosine computation by a factor of l to 4 magnitudes,  and parallel computing.    5. To evaluate the effectiveness of LLSF on large MEDLINE document sets  and compare with the performance of alternate indexing/retrieval systems. | https://reporter.nih.gov/project-details/2392815 | R29 | 1997 | 0.7 |
| DESCRIPTION (Adapted from applicant's abstract): The proposed EEG  MagicMarker software offers a new methodology for the display and  analysis of digitized EEG records. Segments of similar EEG  activity are clustered together, clearly differentiating  background, paroxysmal activity, and patient state transitions,  e.g., sleep stages. The analysis only considers the content of  the current EEG and requires no thresholds or classification  functions derived from a training set.    A novel user interface allows interactive partitioning of the  hierarchical cluster dendrogram in a method already familiar to  many users. Each node can be expanded or collapsed revealing more  or less detail by clicking on the plus or minus sign to the left  of the node. The user manipulates the tree to display the  appropriate partition and prints the report for a summary of their  findings. Each node contains a complete visual summary of the  segments in that power spectrum or by contour plots of delta,  theta, alpha, and beta activities. Hyper-links from the cluster  nodes to the applicant's Insight EEG review software offers  immediate access to pages of interest. | https://reporter.nih.gov/project-details/2034824 | R43 | 1997 | 0.7 |
| DESCRIPTION (Adapted from the Investigator's Abstract): To date no unitary  theory of cortical function has emerged despite a long history of cortical  research. Single cell approaches in primary visual cortex, as exemplified  by Hubeland Wiesel's studies and by recent work on parallel visual pathways,  have produced functional circuit diagrams arguing for hierarchical,  feedforward processing. Alternatively, artificial neural network research  argues that the cortex might represent a distributed feedback circuit in  which intrinsic dynamics converge in stable states that represent  computational solutions. these two types of models predict very different  activation patterns of the circuit. The goal of the research is to  elucidate the three-dimensional spatio-temporal activity patterns intrinsic  to the cortical microcircuit and to identify their underlying circuits.  Studies will be carried out with brain slices from mouse visual cortex using  calcium imaging with a cooled CCD camera, a photodiode array and two-photon  microscope. These techniques allow the investigators to follow the activity  of neuronal ensembles across the entire slice with single-cell and  submillisecond resolution. Specifically, the investigators will (i)  determine the three-dimensional activity patterns present in a brain slice  (ii) establish the anatomical and functional connectivity underlying these  dynamics and (iii) identify neurons playing key roles and study their effect  in altering circuit dynamics. These studies may help determine whether  cortical neurons can activate in preferential labeled lines, as predicted by  feedforward models or in a widely distributed pattern, as predicted by  feedback models, shedding light on the functional units of cortical  microcircuitry and their co-ordination in cortical function as a whole.  Finally, they will help understand the central pathophysiological  consequences of amblyopia and strabismus, as well as help design therapeutic  strategies aimed at compensating for these defect. A more complete  understanding of the circuitry will also improve the analysis of visual  evoked potentials (VEP) and thus the measurement of acuity, contrast  sensitivity and chromatic sensitivity of preverbal children and in early  diagnosis of visual pathologies. | https://reporter.nih.gov/project-details/2485367 | R01 | 1998 | 0.7 |
| A prototype robotic instrument, "Stericulture," has been developed to feed  and harvest cell cultures in an aseptic environment. The instrument is  microprocessor controlled. It was designed to protect the cultures from  contamination and the technologist from hazardous exposure, by  dramatically reducing the handling of the petri dishes. During operation  the repetitive exacting tasks of media dispensing and removal, and petri  dish manipulation are mechanically accomplished within a hood or P3  containment facility. Culture dish lids are removed and replaced, media  is added and removed in timed sequences. The concept can be expanded to  virtually any cell culture lab. The specific aims of this application are  (1) Test the mechanical reliability, accuracy and durability of  Stericulture and its peripheral pumping systems. (2) Conduct beta site  tests at two cytogenetic laboratories and one molecular biology cell  culture laboratory to: a) Evaluate existing computer programming and  written documentation. b) Test the device for reliability in the feeding  and harvesting of cell cultures from amniocytes and biopsies  (fibroblasts). The long-term objective (Phase II) is to develop an  integrated liquid handling system (LHS) with Stericulture as its core.    PROPOSED COMMERCIAL APPLICATION: Development of robotic instrument that  functions in an aseptic environment in a cell culture lab. | https://reporter.nih.gov/project-details/2672772 | R44 | 1998 | 0.58 |
| Colorectal carcinoma is the second leading cause of cancer deaths in the  United States today. In an effort to reduce mortality, Congress  recently included a provision in the Balanced Budget Act of 1997 to  support screening colonoscopy as a means for early detection and removal  of colorectal polyps, the precursors to cancer. In this country alone,  more than 68 million people are eligible for colorectal screening, but  the majority are unlikely to comply with screening recommendations  because of the costs, risks, discomfort, and inconvenience associated  with traditional endoscopy. Furthermore, even if a small fraction of  eligible persons are examined, the number of available  gastroenterologists would be insufficient to perform so many procedures.    We have developed a new technique, called virtual colonoscopy (VC), as  an alternative to screening diagnostic colonoscopy (DC). The procedure  consists of cleansing a patient's colon, inflating the colon with air,  scanning the abdomen with helical computed tomography (CT), and  generating a rapid sequence of three-dimensional (3D) images of the  colon by means of virtual reality computer technology. Although VC  makes possible the visualization of 3D images of the colon in a manner  similar to that of DC, a correct diagnosis depends upon a physician's  ability to identify small and sometimes subtle polyps within hundreds  of 3D images. The absence of visual cues that normally occur with DC  makes VC interpretation tedious and susceptible to error.    With support from a National Science Foundation (NSF) grant, we have  developed a computer-assisted polyp detection (CAPD) system that  calculates areas of abnormal colon wall thickness in helical CT image  data in order to highlight potential polyps in the 3D images. A  physician ultimately determines if each detected lesion represents a  true abnormality. Although we have found CAPD to be sensitive for  finding subtle abnormalities, poor specificity can be attributed to  several obstacles, including imprecise image segmentation, limited  feature analysis, and suboptimal bowel preparation prior to helical CT  scanning. With these challenges in mind, we propose research to perfect  CAPD. Our specific aims are as follows: 1. To develop an image  segmentation algorithm that accurately isolates the colon from helical  CT image data; 2. To improve our polyp detection algorithm with expanded  feature analysis and artificial intelligence methods; 3. To optimize  bowel preparation with digital subtraction of opacified feces and  controlled gas distention; and 4. To validate the accuracy of VC, with  the modifications achieved in the stated aims, by comparing the results  of VC and DC in 200 patients undergoing usual-care colonoscopy.    If VC with CAPD proves accurate and efficient in the diagnosis of  colorectal polyps, it could evolve into a simple laboratory test,  thereby meeting the demand for worldwide colorectal cancer screening. | https://reporter.nih.gov/project-details/2849551 | R01 | 1999 | 0.58 |

**Cluster 7:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| Knowledge discovery and Data Mining (Knowledge Discovery in  Databases-KDD) is a new research field which incorporates  methodologies from artificial intelligence, data bases, and statistics to  address the problem of discovering novel, interesting and useful  patterns (knowledge) hidden in large databases. A prototype KDD  surveillance system for epidemiology named Hawkeye has been  developed at UAB. The goal of this research project is to further  develop Hawkeye into a useful, general-purpose KDD surveillance  system for epidemiology. This will be accomplished in two ways. First,  experiments with the prototype system will be conducted using hospital  infection control data and public health data (CDC). These  experiments will enlist infection control experts, and epidemiologists.  Objective results and subjective feedback will be obtained. These  application-based experiments will further efforts in addressing  fundamental research issues. Second, specific methods improving the  presentation of discovered knowledge to the user will be defined and  implemented. These methods make use of a novel idea called a  phenomenon cluster. | https://reporter.nih.gov/project-details/2411189 | F37 | 1997 | 0.8 |
| Investigating individual neurons has proven very useful in understanding  their functionality. However, all biological systems are composed of a  multitude of neurons that interact-with a variety of sensory and skeletal  muscular systems. Researchers are anxious to have the ability to record  a large number of neuron firing patterns to study how they communicate  and interact in a biological system. Tools such as the multi-channel  electrode have been created to study the interactions of large numbers  of neurons. However, in order to record all the channels, the researcher  must gang several recorders together. This leads to difficulties in  accurately matching all the channels during playback, because of the  complexity of ganged multiple recorders. R.C - Electronics Inc. proposes  to build a low cost, high speed digital recorder utilizing PC technology.  This digital recorder will be able to record continuously to a mass  storage device at an aggregate rate of 1 MHz. It will also provide the  researcher with the ability to acquire up to 96 channels, with each  channel sampled at 10 KHz; higher speeds will be possible with fewer  channels.    PROPOSED COMMERCIAL APPLICATION: There is a market demand for the  digital recorder in the multitude of life science applications engaged  in complex biological system studies. We also believe that this will  be a good industrial recorder, capable of replacing the existing analog  and DAT recorders for acoustic, noise. and vibration studies. There will  also be a market to replace some of the high-end work stations based on  HP, Concurrent and Sun data acquisition systems. | https://reporter.nih.gov/project-details/2445836 | R44 | 1997 | 0.8 |
| DESCRIPTION (Taken from application abstract): In the rapidly changing  health care environment the increasing prominence of managed care has  prompted a greater reliance on formal clinical guidelines to suppose  clinical decisions. Guidelines have been advocated with increasing  frequency to reduce inappropriate care, control geographic variations in  practice patterns and make more effective use of health care resources.  However, guidelines often have little impact on clinical practice because  physicians are unaware of them, lack confidence in them because the  justification for their recommendations is not clear or because they are  inaccessible at the time of patient care or difficult to apply. Guidelines  also quickly become out of date as new research data becomes available.    In order to enhance the quality and usefulness of clinical guidelines, the  general goal of this project is to develop, deploy and evaluate interactive  computer based guidelines, supported by an integrated decision theoretic  model and a linked knowledge base. This arrangement will use patient  characteristics to tailor guideline advice to the individual patient. The  integrated decision model will provide recommendations for situations not  addressed by the guideline and also will help to justify guideline  recommendations by calculating the effectiveness and costs of various  strategies. The proposed system also contains links to a knowledge base  containing the sources of data used in the guideline and the model, so that  physicians using the system can examine the studies supporting the guideline  recommendations. Natural language explanations, generated automatically  based on the structure of the decision model and the guideline will justify  the recommendations. A query capability will enable physicians to look up  specific data from the knowledge base. The linked knowledge based also  ensures that the supported decision model and guideline will be updated  automatically as new research data is published.    The computer system will be used to implement an interactive version of the  Guidelines for Medical Treatment for Stroke Prevention, developed by the  American College of Physicians. The system will be bench tested using a  series of cases abstracted from the General Medicine practice at the Robert  Wood Johnson Medical School. Faculty and house staff internists will serve  as research subjects to perform a field trial of the system. This will  include the extent of previous compliance with the guideline, pre and  post-testing of medical knowledge pertinent to the guideline, and the degree  to which the computer-based guideline changes behavior compared to the  traditional guideline format. A decision theoretic measure of potential  benefit will be calculated by comparing decision model evaluations of  physicians' unaided choices with those recommended by the guideline system. | https://reporter.nih.gov/project-details/2460263 | R01 | 1998 | 0.8 |
| Stuttering is a disorder of speech with a prevalence estimated to be 1 %  of the world's population of school-age children. It is often a  significant communicative problem for the individual, limiting educational  and employment opportunities and social and psychological adjustment. The  etiology of stuttering is unknown, and standardized, successful treatments  for stuttering have not been developed. A major impediment to  understanding the etiology of stuttering and to the development of  successful therapeutic techniques is the lack of understanding of the  physiological bases of the disorder. Stuttering manifests itself as a  breakdown in speech motor processes. The complex variables known to affect  the occurrence of stuttering, such as emotional state or linguistic  complexity, must ultimately have an effect on the physiological events  necessary for the production of speech. Therefore, to understand  stuttering it is essential to understand the physiological mechanisms  underlying disruptions of speech motor processes in stuttering.    The research proposed in the present application addresses this general  question: What is the nature of the movement disorder associated with  stuttering? The specific aims are (1) to determine whether motor processes  show evidence of continuous, underlying disturbances in stutterers'  speech, (2) to assess whether failures in speech movement control in  stuttering are related to autonomic nervous system activity and/or to  metabolic respiratory control, (3) to develop new metrics for the analysis  of physiological signals related to speech and to apply these new metrics  to the assessment of stuttering, and (4) to develop pattern recognition  algorithms to determine if there is a consistent set of physiological  events associated with stuttering. The results of the proposed studies and  those completed in the past years of this project should help us to  understand the complex human behavior that is stuttering. In addition,  work on this project has significant implications for the study of normal  speech production and a variety of motor speech disorders that occur in  neurologically impaired individuals. | https://reporter.nih.gov/project-details/2125802 | R01 | 1995 | 0.41 |
|  |  |  |  |  |

**Cluster 8:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The prediction of the three dimensional structure of a globular protein  from its amino acid sequence along with the mechanism by which protein  folding occurs are among the most important unsolved problems of  contemporary molecular biology. The overall objectives of this proposal  are the continued development and refinement of algorithms which not only  can predict protein tertiary structure using only sequence information as  input but also may provide insights into the folding pathway. To achieve  these goals, this proposal focuses on the lattice based aspects of a  hierarchical approach to protein folding. High resolution lattice models  of proteins, comprised of an alpha-carbon plus reduced off lattice, side  chain description, will provide the overall folding pathways and folded  conformations. The resulting folded lattice structures are estimated for  the alpha-carbons to have a 2-4 angstroms rms deviation from the native  state. Turning to the folding pathways, the predicted molten globule  states and their free energy landscape will be characterized in detail.  The factors responsible for side chain fixation on passage from the molten  globule to the native state will be explored, with particular attention  focused on the interplay of protein sequence and side chain packing.  Specifically this proposal will address the following. (1). A new high  coordination lattice model of proteins will be refined, different side  chain realizations will be examined and the dynamic Monte Carlo algorithms  parallelized. (2). Better empirical free energy functions will be  developed. These include better methods for predicting the propensities  for secondary structure and generalization of the hydrogen bond scheme to  include backbone-side chain hydrogen bonds. To help eliminate misfolded  structures, additional very robust knowledge based rules, such as the  connections in supersecondary structural elements do not cross, will be  included in the interaction scheme. Sequence specific tertiary  interactions including a local burial turn, pair interactions and  generalized cooperative multibody side chain contact templates will be  self consistently derived in the presence of predicted secondary structure  propensities. Then, a recently developed neural network which can  recognize whether 7 by 7 subfragments of sidechain contact maps are  protein like or not will be extended to include sequence specific  preferences for subsequences to adopt specific patterns. This information  will be obtained from a neural network trained on both homologous and non  homologous subsequences that adopt these patterns. Thus, it should be  general and not simply applicable to homologous sequence fragments. (3).  The folding of representative motifs of globular proteins will be  undertaken. Included are the helical proteins such as cytochrome c, whose  predicted folding pathway will be compared to experiment, myohemerythrin,  myoglobin and complement factor, 1c5a. The mixed motif proteins include  ubiquitin, flavodoxin and PRA isomerase, and the beta-proteins include the  16th complement control protein of factor H, 1hcc, alpha-amylase,  plastocyanin and retinol binding protein. (4). To validate the  methodology, additional blind predictions of proteins whose structures are  unknown will be undertaken. Likely candidates include rusticyanin and  erythropoietin. | https://reporter.nih.gov/project-details/2178769 | R01 | 1995 | 0.58 |
| DESCRIPTION (Taken from application abstract): The long term goal of the  work described in this proposal is to develop on-line combined spatial and  symbolic methods for representing, storing, retrieving and visualizing  anatomical information, both as a means for understanding human biological  structure, and as a visual gateway into the rapidly increasing array of  on-line text-based information sources in biomedicine. In this proposal we  will address some of the fundamental problems involved in combining spatial  and symbolic anatomic information, and will test solutions to these problems  in a World Wide Web based 3-D anatomy information system for the thoracic  viscera. Clinicians, researchers and students will be able to use this  system to retrieve specific anatomic knowledge, in the form of  dynamically-generated interactive 3-D scenes and corresponding symbolic  information, without the need to consult hard copy atlases or to navigate  through irrelevant computer-based images before finding the needed  information. In order to build this system we will need to address  fundamental problems in spatial modeling, organization of these models in a  knowledge based spatial database system, and access to the models via an  on-line user interface and spatial query processor. The specific aims for  this proposal are: 1) develop a knowledge base that organizes and  integrates spatial and symbolic models of anatomy, 2) implement an anatomy  information system that combines knowledge-based spatial and symbolic  retrieval with dynamically generated 3-D scenes, 3) develop methods for  smoothly rendering and interacting with the scene in real-time, and 4)  evaluate the system by integrating other spatial data and by providing it to  anatomy medical students and radiation treatment planners. Accomplishment  of these aims will lead to a useful information system for the anatomy of  the thoracic viscera that can be enhanced with new technology such as high  performance graphics and virtual reality. Moreover, the information  framework and methods we establish in this project will be generalizable not  only to gross anatomy in different regions of the body and to different  anatomical databases, but also to the management of structural information  pertaining to cellular and molecular biology, as well as developmental and  neurobiology. | https://reporter.nih.gov/project-details/2032431 | R01 | 1997 | 0.58 |
| Ordered categorical variables arise frequently in cancer clinical trials  and other biomedical studies. The statistical procedures for analyzing  such data are well known and software for performing the analysis is  readily available. The basic idea is to condition on the margins of the  contingency table created by the categorical data and thereby obtain a  distribution free test that automatically corrects for ties. Despite the  popularity of this conditional approach for analyzing ordered categorical  data there has been very little work done on power and sample-size  considerations at the design phase. A biomedical investigator about to  launch a clinical trial for comparing two treatments with ordered  categorical outcomes will find it extremely difficult to determine what  sample size is needed. Either the investigator must assume that the data  are continuous, or else that the data are binary, since these are the only  cases for which reliable methods and software are available. Both  approaches are inappropriate for ordered categorical data. We propose to  fill the void by providing new exact and Monte Carlo methods that provide  accurate power and sample-size estimates for conditional tests on ordered  categorical data. | https://reporter.nih.gov/project-details/2434545 | R44 | 1997 | 0.58 |
| There is growing recognition that, while practice guidelines can  predispose physicians to behavioral change, even the most respected and  clinically relevant guidelines often do not affect actual practice.  Despite the apparent reluctance of physicians to incorporate published  practice guidelines into their clinical routine, there are signs that  immediate automated reminders can have a profound influence on physician  behavior, and on the cost and quality of medical care. As practice  guidelines proliferate and become more complex, the dilemma facing those  who design practice guidelines is the tradeoff between the  comprehensiveness of the guidelines and the likelihood they will be  retained and implemented by the physicians for whom they are intended.  This same complexity can complicate computer-based implementation of  guidelines. For example, the dependence of guidelines on patient  preferences and individual risk factors makes the rule-based approach  impractical for many guidelines.    Intelligent decision systems (IDS) are specifically designed to implement  complex, flexible guidelines tailored to varying clinical circumstances.  These systems provide advice based on a tailored decision-analytic model.  But these systems assume that a physician-user will be persuaded by a  review of the decision model and its associated quantitative results.    The goal of this work is to develop and test improved explanation methods  for quantitative decision models so that intelligent decision systems can  be used in a consultative mode. We have developed a program called QxQ,  which uses symbolic reasoning to provide qualitative text explanations for  the results of decision trees. This work is intended to extract the key  architecture- and domain-independent elements of QxQ's methodology, and to  reimplement and test them as part of MIDAS, an existing IDS. First, we  will develop a domain- and system-independent version of QxQ, called EQxQ  (for essential QxQ), by applying it to decision models constructed in  MIDAS. In collaboration with the developers of MIDAS, we will identify the  key independent elements necessary to generate explanations from an IDS.  Second, we will devise and implement additional explanation methods in  EQxQ to allow explanation of complex modeling constructs, such as Markov  models, cycle trees, and cost-effectiveness models. Third, we will test  the robustness of the MIDAS-EQxQ system by using it to implement a  decision model underlying a clinically relevant practice guideline.  Finally, we will establish and pilot test the routine use of the program  in an outpatient clinic. We will survey the physician-users to determine  the strengths and weaknesses of the program. Using pilot data from the  program's routine use, we will design a prospective clinical evaluation of  the system and its explanations by measuring their effect on patient  outcome. Our study will provide data regarding the clinical utility of  intelligent decision systems. | https://reporter.nih.gov/project-details/2714217 | R29 | 1998 | 0.32 |
| DESCRIPTION (Adapted from applicant's abstract): Visuomotor integration  depends on a remarkable coherence among a number of interrelated  subprocesses such as pattern recognition, pattern discrimination, decision  to move, and guidance of movement. The brain is able to integrate these  elementary cognitive processes by coordinating the activities of diverse  neural structures in the face of continuously varying processing demands.  The question of how this coordination operates is central to understanding  the neural basis of visuomotor function. This proposal aims to develop new  analytical tools to investigate the coordinated activity of distributed  neuronal ensembles in the cerebral cortex of humans and non-human primates  performing simple visuomotor tasks. It is motivated by recent theoretical  developments (Bressler 1994, 1995, 1996, and 1997) predicting a general  cortical mechanism allowing the flexible large-scale functional coordination  of interacting neuronal ensembles. Hypotheses concerning this mechanism  will be tested by analysis of field potential data recorded at NIMH from  macaque monkeys performing a visuomotor pattern recognition task. The  challenge is to develop and test new analytic approaches that characterize  the multiple, complex interactions of large-scale distributed cortical  networks. Earlier analysis of a small portion of this NIMH data set,  reported in Nature in 1993, revealed shifting patterns of multi-site  cortical synchronization during visuomotor processing, and implicated  synchronization in the formation of functional relations within and between  cortical areas. Standard pairwise techniques were employed to measure  synchronization between field potential signals. Here, novel methods of  time-series analysis are proposed that go beyond the simple detection of  network interactions. Advances in signal processing technology will be  utilized to also derive multi-site interaction patterns, to analyze the  dependencies of functional relations on particular groups of neurons, and to  measure the flow of information between cortical regions. This  collaborative project will draw on the complementary strengths of Drs.  Bressler and Ding. Dr. Bressler brings to the project over 15 years of  experience in cognitive neuroscience, with expertise in the recording and  analysis of neuroelectric data from humans to animals. He will provide  theoretical oversight and the application of analytic tools to the field  potential data set. Dr. Ding, although relatively new to cognitive  neuroscience, has over 10 years of experience in linear and nonlinear  dynamical systems analysis. He will provide the development of new  analytical methods from a comprehensive dynamical systems perspective. This  work is expected to (1) produce new insights into the dynamics of cortical  information flow in visual perception and motor performance, (2) make  available new digital signal processing tools for the investigation of large  scale neural systems underlying other cognitive functions and, (3) provide a  fresh perspective on the design of complex architectures for the execution  of cognitive tasks by artificial neural network systems. | https://reporter.nih.gov/project-details/2891078 | R03 | 1999 | 0.32 |

**Cluster 9:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| DESCRIPTION: The overall objective of the proposed research is to  determine the mechanisms responsible for the thermally induced changes  that occur in the coefficients of thermal expansion of dental porcelains  as a result of multiple firings, slow cooling and post-soldering  operations. These changes in porcelain thermal expansion are  detrimental and produce cracks in porcelain fused to metal restorations  (PFM) either shortly after firing or sometime later. Either event  causes additional costs and possibly additional trauma to the patient.  There are six Specific Aims. The first is to explain the thermal  expansion behavior of porcelains on the basis of leucite and sanidine  volume fractions, microcrack densities, an leucite particle surface  areas. The second is to determine the effect of localized cooling rate  differences on the thermal instability of dental porcelain during  multiple firing, cooling, and isothermal anneals. The third is to  determine the role of H2O as a glass modifier in dental porcelain and  specifically in the crystallization of leucite and sanidine. The fourth  is to develop firing schedules for various porcelain-metal systems to  minimize the thermally induced microstructural changes that lead to  thermal expansion changes.The fifth is to modify porcelains to render  them more resistant to thermal expansion changes. The sixth to develop  a prototype expert system computer program for solving thermal expansion  mismatch problems. Techniques used in pursuit of these aims include  quantitative X-Ray diffraction, scanning electron microscopy with energy  dispersive X-ray spectroscopy, quantitative stereology, and conventional  and laser dilatometry. | https://reporter.nih.gov/project-details/2129906 | R01 | 1995 | 0.85 |
| DESCRIPTION:(Adapted from the Investigator's Abstract) The goal of this  grant is to determine if and when cancer risks can be estimated by  establishing record-linkages between statewide cancer surveillance systems  and occupational cohorts. More specifically, the aims of this study are to:  (1) determine the feasibility of utilizing statewide cancer surveillance  systems in the evaluation of cancer incidence within occupational cohorts;  (2) compare and contrast the relative merits of standardized incidence  ratios (SIR) with standardized mortality ratios (SMR) as determined from  cancer surveillance systems incidence data and death certificate mortality  data, respectively; and (3) provide recommendations concerning how and when  statewide cancer surveillance systems should be utilized in the evaluation  of occupational cohorts.    SMR and SIR estimates will be calculated and compared for three occupational  cohorts (e.g., Highway Maintenance, 3M, Conwed). SMR analyses have already  been completed for the Highway Maintenance and 3M cohorts; a mortality  update and SMR analysis will be conducted for the Conwed cohort (1988-1995).  Cancer morbidity information, for the SIR analysis, will be determined by  linking the three cohorts with the Minnesota Cancer Surveillance System  (MCSS). Residency status will be required before person-years can be  calculated, however, because inclusion in the MCSS is restricted to  Minnesota residents. Linkages to other data sets will be used to determine  the Minnesota residency status. Sensitivity analyses will be used to  evaluate confounding and follow-up bias. Standardized mortality ratios will  be compared to standardized incidence ratios for the Highway Maintenance,  3M, and Conwed cohorts. Finally, this study will evaluate the utility and  limitations of cancer surveillance systems as a tool for occupational cancer  research; recommendations for its use will be developed. | https://reporter.nih.gov/project-details/2039440 | K01 | 1997 | 0.85 |
| Asthma is a highly prevalent disease with a significant morbidity leading  to emergency room visits, hospitalizations and absenteeism from work and  school. Two major factors contributing to asthma morbidity (and mortality)  are failure to detect asthma decompensation and use of inappropriate  treatment strategies. Asthma severity can be assessed by clinical symptoms  and peak flow measurements. There are internationally accepted guidelines  for asthma management. However, rapid translation of asthma symptoms into  a change in therapy is frequently impeded by the patient's inability to  understand self management, reluctance to contact the health care  provider, or delays in feedback from the provider to the patient. The  objective of this proposal is to develop a prototype Automated System for  Telephonic Home Management of Asthma (ASTHMA) in analogy to an existing  and successful on-line management system for diabetes.    Specifically, we will transform the algorithms for asthma management into  a format suitable for computer implementation, develop a computer platform  for the testing of these algorithms, and conduct off-line clinical tests  to establish the technical merit and feasibility of a prototype system.  This will be accomplished in 200-250 ambulatory children and adults with  asthma who are served in three clinics at the University of Miami School  of Medicine. A comparison between the "mock recommendations" of the  prototype system with the physicians' actual interventions will test the  feasibility of the computerized system. The goals of the ASTHMA system  described in this proposal are to save time (efficiency), decrease  morbidity (efficacy), and increase patient and health care provider  satisfaction (attractiveness).    PROPOSED COMMERCIAL APPLICATION:  Computer-assisted home management of asthma offers a significant  commercial opportunity as systems for this purpose will potentially be  needed in all major asthma care centers. | https://reporter.nih.gov/project-details/2421250 | R43 | 1997 | 0.85 |
| The amino acid sequence of a protein uniquely determines its tertiary  structure. Deciphering this relationship, the protein folding problem has  become increasingly important to molecular biologists. DNA sequencing has  become routine, but structural experiments remain very difficult.  Computational strategies are needed to help address this problem.    This proposal describes a strategy to identify the location of alpha-  helices and beta-strands throughout the sequence. A method for using off-  lattice simulations of a polypeptide chain to identify secondary structure  preferences in the ensemble average is proposed. Once secondary structure  is located, computational methods exist for generating plausible tertiary  structures. However, these combinatorial strategies give rise to a large  number of alternative structures which are difficult to distinguish from  the correct fold. Experimental and theoretical methods for clarifying the  distinction between correctly folded structures and their misfolded  counterparts will be considered.    In a new direction, we propose to develop a multiple sequence analysis  strategy to relate sequence and structure to function. In particular, we  will focus on identifying the binding sites on the G-alpha family of  GTPases for the relevant G-protein coupled receptors, G-beta-gamma and  downstream effectors. We plan to continue to develop a genetic algorithm  for the construction of polypeptide loops subject to a series of  constraints. This method will be used to model the loop regions of G-  protein coupled receptors involved in the interaction with peptide ligands  and the hetero-trimeric G-protein complex.    Finally, we propose to develop a new method to compare structures based on  the area of the minimal "soap film" that could join them following the  appropriate rotation and translation of one structure relative to another.  This provides a natural way to circumvent the gap penalty problem that  plagues current structure alignment algorithms. | https://reporter.nih.gov/project-details/2608876 | R01 | 1998 | 0.77 |
| DESCRIPTION: (Applicant's Abstract)    The goal of this five year research is to provide new knowledge on: 1) the  causes of drug use and abuse in inner city African-American and White  adolescents, and 2) the most cost-effective and cost-beneficial  family-focused prevention interventions. The research aims of this  randomized clinical trial involving over 800 at-risk Washington, D.C.-area  families of 7 to 11 year olds in components of the 16 week Strengthening  Families Program (SFP) are: 1) to test the differential efficacy of the  three major components of SFP-parent training (PT), children's skills  training (CT), and family skills training (FT) compared to a minimal contact  control (MT) on precursors and substance use/abuse employing a 2 x 4  experimental design; 2) to estimate the cost-benefit and cost-effectiveness  of the SFP components with 400 African American families and 400 White  families; and 3) to empirically test several competing etiological and  intervention theories of drug use/abuse underlying the intervention design  by analyzing the survey research data from 1,600 12 to 19 year olds and the  800 intervention families using structural equation modeling, mediational  analysis, hierarchical linear modeling, latent growth modeling, and  transactional modeling (cross-lagged SEM, artificial intelligence, and  neural network simulations) for gender, race, and other sub-groups. The  rationale for the component design is that a major question exists  concerning the efficacy of child-only interventions that group high-risk  children together as well as the added benefit of working with the family  together compared to parent-training only. Demonstrating the importance of  involving parents and working with families would be very valuable  information for prevention providers. Additional co-variate outcome  analyses will be conducted to answer practical questions conceming the  efficacy of the four interventions for different subgroups of clients. A  strong process evaluation including a management information system (MIS)  conducted by University of Maryland researchers wili document implementation  process and provide feedback to improve implementation fidelity. Additional  health services sub-aim questions will be addressed through analyses of  recruitment, attrition, consumer satisfaction, participation variables, and  client and trainer characteristics linked to outcomes. Seven nationally  recognized consultants have agreed to support the culturally-specific  program development, implementation, and data analyses. | https://reporter.nih.gov/project-details/2882620 | R01 | 1999 | 0.77 |

**Cluster 10:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| DESCRIPTION (Adapted from the Investigator's Abstract): Numerous  restriction mapping programs have been devised since 1978, but none of them  is as precise as manual map construction using digest gel photos, pencil,  and graph paper. Given realistic data, programs usually find hundreds or  even thousands of solutions, only one of which can be correct. Because of  this, slow and demanding manual techniques are still in common use. The  researchers propose to develop public domain software implementing a  complete restriction mapping environment which will be far more powerful  and useful than current restriction mapping software. This system will:  1) find solutions more quickly than any existing software; 2) find far  fewer false maps than any existing software; 3) allow the user to "steer"  the entire mapping process (if desired); and 4) guide the user with  detailed, expert advice on handling specific mapping problems as they  arise. In addition, the system will have: 5) built-in extensibility, so  that simple modifications will also allow other genetic marker mapping  problems to be solved by the same program and 6) the ability to record all  user activity transparently, providing quantitative data on successful  mapping strategies.    The system will be based on extant restriction mapping programs, but it  will overcome their limitations by including the following additional  capabilities: 1) All known heuristics will be implemented. (Heuristics  are logical rules of thumb which guide the search towards a true solution.)  Existing programs use only a small subset of the known heuristics. Every  added heuristic will speed up the search and reduce the set of possible  solutions to a problem. 2) Hand mapping will be simulated on-screen via a  "what you see is what you get" graphical user interface, with users  choosing fully automatic mapping (the default), fully manual mapping (with  simulated pencils and log paper), or various in-between levels of  semi-automatic assisted mapping. This will allow users to control or  adjust any part of the mapping process if they so desire. 3) An expert  system (a program which can answer queries and make decisions by consulting  a knowledge database) will guide users through the mapping process. It  will assist the process of data acquisition, help the user solve  difficulties, and tutor inexperienced users.    Considerable preliminary design work has already taken place, so  implementation can begin almost at once. Software will be developed and  tested iteratively ("rapid prototyping") to assure end-user satisfaction.  To insure portability at the source code level, the two major modules of  the system (fragment length derivation and mapping) will be coded in C++  using Boochs' object-oriented design methodology, and the user interface  will be designed using a portable interface-building tool that works on a  variety of computing platforms (including DOS machines, and the Macintosh).  The expert system will be implemented using CLIPS, a portable C-based  public domain expert system shell. | https://reporter.nih.gov/project-details/2209202 | R01 | 1995 | 1 |
| This proposal presents a set of principles of human information processing  that have been developed by the investigator and describes a program of  collaborative research designed to assess the adequacy of these principles  for modeling cognitive processes in normal and disordered individuals. In  brief, the principles state the human cognition occurs through a graded,  stochastic, adaptive, interactive and distributed process. The ultimate  goal is the further development and understanding of the set of principles  and a further exploration of the implications of the principles for normal  and disordered cognition. The principles serve as the focus of specific  research projects on the following topics: general laws and regularities  of information processing; the role of context in visual information  processing; the mechanisms of attention and neuromodulatory deficits of  attention; and the mechanisms of attention and neuromodulatory deficits of  attention; and the mechanisms of accessing representations of the sounds  and meanings of printed words. Each project involves the comparison of  the results of computer simulation models based on the principles to the  results of psychological experiments designed to produce evidence relevant  to the assessment of the adequacy of the principles. The proposal also  describes plans for the further incorporation of principles emerging from  neuroscience and for the dissemination of concepts and simulation tools  that are crucial to facilitate exploration of the principles by other  researchers. | https://reporter.nih.gov/project-details/2239797 | K05 | 1995 | 1 |
| DESCRIPTION (Adapted from applicant's abstract): This K02 application  out-lines research and career development plans to investigate the role of  sleep regulation in the developmental psychobiology of affective disorders.  The first goal is to extend a well-established line of research based on  measures of EEG sleep and cortisol in child and adolescent depression by  investigating: (1) abnormalities in sleep and cortisol regulation focusing  on the sleep-onset transition; (2) normal maturational changes in sleep and  cortisol regulation relevant to these abnormalities; and (3) the predictive  validity of sleep and cortisol abnormalities in longitudinal clinical  follow-up.    The second goal is to further develop and investigate a larger developmental  model of sleep regulation. The model emphasizes close links between sleep  regulation and neurobehavioral systems involved in the regulation of affect  and arousal which are modulated in regions of prefrontal cortex (PFC).  Matur-ational changes in PFC-subcortical circuits influencing sleep,  arousal, and affect are hypothesized to contribute to sleep changes  associated with depression, particularly near the transition from  wakefulness into sleep. Based on this model, predictions are made regarding  sleep changes in the development of affective disorders and effects of sleep  deprivation of affective regulation. The long-term goal of this work is to  understand mechanisms of dysregulation which may lead to more effective  treatment of early onset affective disorders.    Career development activities to support these goals are described. Further  advancing this line of investigation will require increased knowledge and  skills in four areas: the development of affect regulation; the relevant  neurocircuitry and its development; more direct measures of the neural  systems of interest; and statistics. The candidate will pursue these goals  through course work, focused readings, and supervised learning experiences  and collab-orations with scientists locally and nationally. | https://reporter.nih.gov/project-details/2415800 | K02 | 1997 | 1 |
| DESCRIPTION: Numerous practical and theoretical problems could be addressed  if we had a better understanding of the auditory mechanisms underlying  phonetic recognition. Among the practical applications of this knowledge  are: (1) the improvement of speech synthesis devices, (2) the development  of robust speech recognition devices, (3) the development of acoustically  based training devices for hearing-impaired speakers, and (4) improvement in  Cochlear-implant signal processors. The proposed experiments fall into  three major categories. One set of experiments follows in a rather direct  way from vowel perception studies conducted during the previous grant  period. These experiments address issues such as the role of dynamic  spectral cues and voice fundamental frequency in vowel perception. A second  series of experiments address more fundamental issues regarding the spectral  representations that control phonetic quality. A major goal of these  experiments is to test the validity of a method of representing speech that  was developed during the previous grant period. The "Masked Peak  Representation" (MPR) was developed as an alternative to both formant  representations and whole spectrum models. The MPR involves a series of  spectral manipulations that are designed to remove aspects of the spectrum  that do not appear to have a strong influence on phonetic quality, while  retaining those features that are most relevant to phonetic quality  judgments. The MPR will be evaluated with: (1) an experiment comparing  MPR-based predictions of perceived phonetic distance with those of a more  traditional auditory model, (2) speech recognition tests that use a Hidden  Markov Model to map sequences of MPR spectra onto words or phonetic  segments, and (3) listening tests with speech resynthesized from MPR  spectra. A third set of studies is aimed at modeling the low-level auditory  mechanisms that are responsible for spectrum analysis. The goal of this  work is to evaluate a model of spectrum analysis that is carried out by the  central auditory system rather than the auditory periphery. A software  simulation of the model will be developed in an effort to determine the  extent to which the central-spectrum model can account for a broad range of  findings from the auditory psychophysics literature. Experiments are also  proposed that address the implications of this model for vowel perception  and for the representation of pitch and periodicity. | https://reporter.nih.gov/project-details/2608271 | R01 | 1998 | 1 |
| DESCRIPTION: The proposed research will focus on the mechanisms of electron  transfer and proton pumping in respiration. Special emphasis will be on  cytochrome oxidase, the primary site of coupling between the two processes.  Our goals include the following: 1) The mechanism of the reduction of  dixoygen to water by cytochrome oxidase will be studied under a variety of  conditions with the flow-flash method, which uses the photolability of the  CO complex to initiate the redox activity with O2. Time-resolved  multichannel optical absorption spectroscopy will be used to follow the  kinetics of electron and proton transfer on time scales of nanoseconds to  milliseconds. Singular value decomposition and global exponential fitting  methods will be applied to analyze the kinetics and determine the UV-Vis  spectra of the transient intermediates. These studies should provide new  insight into the mechanism of the dioxygen reduction by cytochrome oxidase.  2) Alternatives to CO photodissociation will be used to investigate the  reaction of O2 with cytochrome oxidase. The reaction of unliganded reduced  cytochrome oxidase with oxygen will be studied using a superfast direct  mixing method, pulsed-accelerated-flow (PAF). The reaction of oxygen with  the unliganded enzyme will also be investigated using O2 which is produced  in situ on any relevant time scale by photodissociating synthetic dioxygen  carriers such as dicobalt u-peroxo polyaine complexes. Both the PAF method  and the photodissociation of the dicobalt u-peroxo and u-superoxo polyamine  complexes represent new approaches to study the fast dioxygen reactions of  cytochrome oxidase and both avoid the mechanistic ambiguities associated  with the fate of photodissociated CO in transitional flow-flash experiments.  3) The mechanism of the redox-linked proton pump in cytochrome oxidase will  be investigated. The kinetics of electron transfer and proton pumping upon  flash-induced oxidation of cytochrome oxidase reconstituted into  phospholipid vesicles will be monitored using time-resolved optical  absorption spectroscopy. The proton pumping reactions will be probed by pH  indicators located int he extra-vesicular space, trapped inside the  vesicles, or covalently bound to the lipid or protein. These studies will  allow us to correlate proton pumping events with individual steps in the  dioxygen/cytochrome oxidase redox cycle and will provide a foundation for a  structural model of the energy transduction mechanism in cytochrome oxidase. | https://reporter.nih.gov/project-details/2655009 | R01 | 1998 | 1 |

**Cluster 11:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| We will continue our development of methods for recognizing and  representing functional domains in biological sequences. This  includes methods to identify regulatory sites in DNA starting  from unaligned sequences, and to develop models that will allow  new sites to be accurately predicted. This will involve the  adoption of better statistical models so that the most  significant alignments can be more readily obtained. We will  also develop improved methods for recognizing functional motifs  in RNA sequences that are composed of both sequence and  structure. These methods will be useful for identifying  regulatory domains that operate post-transcriptionally, and also  for determining the common motifs in RNAs selected in vitro for  particular activities. And we will further enhance methods for  representing conserved domains in protein families that new  members of the families can be identified more reliably. This  will involve the use of neural network methods that optimize the  discrimination of protein family members from other sequences in  the database that are not members of the family.    We will also continue several collaborations with biologists who  can take advantage of our methods in their work, and develop new  collaborations as opportunities arise. | https://reporter.nih.gov/project-details/2394730 | R01 | 1997 | 0.88 |
| Telemammography requires high spatial resolution, the transfer of large  image files and rapid lossless transmission of images. Current  teleradiology techniques fail to meet these requirements. A novel  approach addressing these difficulties, involving robotic control of  image acquisition, optical magnification and unique transmission  strategies licensed from BellSouth Telecommunications, is proposed.  Phase I goals include prototype development, software design and  testing, objective characterization of the system and subjective system  evaluation by a board certified radiologist. Following completion of  Phase I, a Phase II application will be filed. Phase II will include  optimization of system configuration, extensive testing and evaluation,  both in-house and by consulting radiologists, culminating in clinical  trials at a minimum of three sites. Following this we expect to file for  FDA clearance of the device. This research should result in affordable,  high resolution telemammography system that will provide better  service to remote locations. With this system, even remote locations  could have access to expertise and rapid diagnosis, affording women  better care in a stressful time. This novel approach will have  application to all of teleradiology, especially those specialties requiring  a high spatial resolution.    PROPOSED COMMERCIAL APPLICATION: The proposed project will result  in a telemammography system that allows re-acquisition of high spatial  resolution images by a remote radiologist via a roboticly controlled X-  ray viewer. This system will allow transmission of high quality  mammographic images from a center to a remote physician or form a  remote hospital or clinic to a center for a more rapid referral and/or  diagnosis. This improved affordable system for the transmission of  radiologic images will have application to teleradiology in general. | https://reporter.nih.gov/project-details/2422962 | R43 | 1997 | 0.88 |
| DESCRIPTION (Adapted from applicant's abstract): Visuomotor integration  depends on a remarkable coherence among a number of interrelated  subprocesses such as pattern recognition, pattern discrimination, decision  to move, and guidance of movement. The brain is able to integrate these  elementary cognitive processes by coordinating the activities of diverse  neural structures in the face of continuously varying processing demands.  The question of how this coordination operates is central to understanding  the neural basis of visuomotor function. This proposal aims to develop new  analytical tools to investigate the coordinated activity of distributed  neuronal ensembles in the cerebral cortex of humans and non-human primates  performing simple visuomotor tasks. It is motivated by recent theoretical  developments (Bressler 1994, 1995, 1996, and 1997) predicting a general  cortical mechanism allowing the flexible large-scale functional coordination  of interacting neuronal ensembles. Hypotheses concerning this mechanism  will be tested by analysis of field potential data recorded at NIMH from  macaque monkeys performing a visuomotor pattern recognition task. The  challenge is to develop and test new analytic approaches that characterize  the multiple, complex interactions of large-scale distributed cortical  networks. Earlier analysis of a small portion of this NIMH data set,  reported in Nature in 1993, revealed shifting patterns of multi-site  cortical synchronization during visuomotor processing, and implicated  synchronization in the formation of functional relations within and between  cortical areas. Standard pairwise techniques were employed to measure  synchronization between field potential signals. Here, novel methods of  time-series analysis are proposed that go beyond the simple detection of  network interactions. Advances in signal processing technology will be  utilized to also derive multi-site interaction patterns, to analyze the  dependencies of functional relations on particular groups of neurons, and to  measure the flow of information between cortical regions. This  collaborative project will draw on the complementary strengths of Drs.  Bressler and Ding. Dr. Bressler brings to the project over 15 years of  experience in cognitive neuroscience, with expertise in the recording and  analysis of neuroelectric data from humans to animals. He will provide  theoretical oversight and the application of analytic tools to the field  potential data set. Dr. Ding, although relatively new to cognitive  neuroscience, has over 10 years of experience in linear and nonlinear  dynamical systems analysis. He will provide the development of new  analytical methods from a comprehensive dynamical systems perspective. This  work is expected to (1) produce new insights into the dynamics of cortical  information flow in visual perception and motor performance, (2) make  available new digital signal processing tools for the investigation of large  scale neural systems underlying other cognitive functions and, (3) provide a  fresh perspective on the design of complex architectures for the execution  of cognitive tasks by artificial neural network systems. | https://reporter.nih.gov/project-details/2563487 | R03 | 1998 | 0.88 |
| DESCRIPTION: Stage matched interventions for stress management that are interactive and individualized, and are delivered proactively to entire populations can have unprecedented impacts. Computer based expert systems linked to self-help manuals can be as effective as counselors but at much lower cost and greater accessibility. Stress is an important cause of cancer and other chronic and acute diseases and is one of the most costly behaviors in terms of health care, job performance and disability. Fifty million Americans do not practice effective stress management. Existing programs are action-oriented and are designed for the 30% of populations who are prepared to take action. Stage matched programs can meet the needs of all; the 45% in the Pre-contemplation stage and the 25% in the Contemplation stage. Phase I of this Fast-Track research will demonstrate the feasibility of recruiting 70% of at-risk populations and the acceptability of the expert system interventions. Phase II will complete recruitment of 1200 participants randomly assigned to treatment or control and can demonstrate efficacy of these interventions over six months. Follow-up over 18 months can show increasing impact long after the intervention. Effective and cost-effective stress management systems can be broadly disseminated with consistent quality and user friendly acceptability. | https://reporter.nih.gov/project-details/2869454 | R44 | 1999 | 0.88 |
| Stuttering is a disorder of speech with a prevalence estimated to be 1 %  of the world's population of school-age children. It is often a  significant communicative problem for the individual, limiting educational  and employment opportunities and social and psychological adjustment. The  etiology of stuttering is unknown, and standardized, successful treatments  for stuttering have not been developed. A major impediment to  understanding the etiology of stuttering and to the development of  successful therapeutic techniques is the lack of understanding of the  physiological bases of the disorder. Stuttering manifests itself as a  breakdown in speech motor processes. The complex variables known to affect  the occurrence of stuttering, such as emotional state or linguistic  complexity, must ultimately have an effect on the physiological events  necessary for the production of speech. Therefore, to understand  stuttering it is essential to understand the physiological mechanisms  underlying disruptions of speech motor processes in stuttering.    The research proposed in the present application addresses this general  question: What is the nature of the movement disorder associated with  stuttering? The specific aims are (1) to determine whether motor processes  show evidence of continuous, underlying disturbances in stutterers'  speech, (2) to assess whether failures in speech movement control in  stuttering are related to autonomic nervous system activity and/or to  metabolic respiratory control, (3) to develop new metrics for the analysis  of physiological signals related to speech and to apply these new metrics  to the assessment of stuttering, and (4) to develop pattern recognition  algorithms to determine if there is a consistent set of physiological  events associated with stuttering. The results of the proposed studies and  those completed in the past years of this project should help us to  understand the complex human behavior that is stuttering. In addition,  work on this project has significant implications for the study of normal  speech production and a variety of motor speech disorders that occur in  neurologically impaired individuals. | https://reporter.nih.gov/project-details/2443589 | R01 | 1997 | 0.76 |

**Cluster 12:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| DESCRIPTION: This application proposes to develop, implement, and evaluate  a world wide web-based computerized decision support system (CDSS) to  facilitate information exchange and guide interactions between  geographically distributed physicians and centrally-located experts in bone  marrow transplant (BMT) follow-up care. The CDSS will include standard  practice guidelines and research findings specific for the long-term  follow-up (LTFU) of patients post-BMT but will be designed to be adaptable  to other disease and treatment situations. Key elements required for the  conduct of the project are already in place, including an ontology of  long-term follow-up, diagnostic pathways, and practice guidelines; a  multidisciplinary team with broad experience; a high volume of follow-up and  consultation demand; and a network of over 2,000 primary specialists caring  for the patients in a wide variety of practice settings. Each year the LTFU  unit receives over 5,000 pieces of patient-care mail, sends 4,000 letters,  returns 8,000 phone calls, and mails over 1,200 protocols, consent forms,  and medical recommendations.    The proposed project will complete and refine a networked CDSS, conduct a  phase II pilot study of clinical use of the CDSS within the bone marrow  transplant center, conduct a phase III randomized clinical trial of the  benefit of the CDSS with over 250 primary care physicians randomized to  either CDSS or the existing method of follow-up, and evaluate the impact of  CDSS on physician behavior and practice efficiency. Endpoints for the phase  III portion of the project include patient outcomes and complications,  quality of life, cost of patient care, physician satisfaction, and frequency  of accessing the protocols/guidelines. An attempt will also be made to  identify factors predicting the success of the CDSS. | https://reporter.nih.gov/project-details/2546251 | R01 | 1997 | 0.093 |
| The goal of the proposed research is to develop computerized radiographic  methods for measuring bone structure for use in quantitatively assessing  osteoporosis and risk of fracture. We will investigate the characteristics  of trabecular bone structure in digital radiographs in the spine, hip and  extremities using computerized texture analyses. We believe that our  methods have the potential to aid in the assessment of osteoporosis and  that the use of both BMD and bone structure information should improve the  predictive value for assessing fracture risk over that obtainable with BMD  alone.    We will create a database in order to quantify the characteristic features  of the trabecular pattern in high-resolution radiographic bone images of  patients with varying degrees of osteoporosis, as well as in normal  subjects. Specifically, we plan to (l) develop computerized texture  analysis schemes for the automatic assessment of bone structure in  digitized bone radiographs, (2) investigate the effects of various  parameters of the image acquisition system, as well as of the analysis  schemes themselves, on performance and (3) evaluate the efficacies of the  computerized schemes in predicting risk of fracture as compared to a  current method of measurement [dual-energy x-ray absorptiometry (DXA)]  using a large clinical database.    Methods that are capable of analyzing bone structure of trabeculae, along  with bone mass measures, are expected to give additional insight to the  evaluation of osteoporosis and risk of fracture. Our scheme is unique in  that it attempts to quantify automatically the risk of fracture from  texture analyses (Fourier analysis, multi-fractal analysis, gradient  analysis and artificial neural networks) of the bone trabecular pattern as  present in high-resolution radiographs of the spine, hip and extremities.  The potential significance of this research project lies in the fact that  if the detection of high-risk patients could be accomplished with a  reliable, low-dose, economical system, then screening for osteoporosis  could be implemented more broadly, thereby allowing earlier treatment and  a reduction in the risk of fracture. | https://reporter.nih.gov/project-details/2390534 | R01 | 1997 | 0.093 |
| DESCRIPTION (Taken from application abstract): We propose to develop  automated techniques to facilitate classification and pattern recognition in  biomedical data sets. These techniques will involve development of novel  neural network architectures, as well as formulation of principles governing  their creation and explanation of results. Specifically, as a solution to  the problem of recognizing infrequent categories, we will develop  hierarchical and sequential systems of feedforward neural networks that make  use of information such as (a) prior knowledge of the domain, and/or (b)  natural clusters defined by clustering or unsupervised learning methods to  develop intermediate classification goals and utilize a divide-and-conquer  approach to complex classification problems. Additionally, we will develop  generic tools for pre-processing input data by making transformations of  original data, reducing dimensionality, and producing training and test sets  suitable for cross-validation and bootstrap. We will build tools for  evaluating results that measure calibration, resolution, importance of  variables, and comparisons between different models. Furthermore, we will  develop standardized interfaces for certain existing classification models.  We will use a component-based architecture to build our neural network and  write interfaces to existing classification models (e.g., regression trees,  logistic regression models) so that they can be interchanged in a  user-friendly manner. We will use our preprocessing modules to prepare data  to be entered in a variety of classification models. The results will be  evaluated in isolation, and later combined to test the hypothesis that the  combined system performs better in real biomedical data sets in terms of  calibration, resolution, and explanatory power.    This research will (a) quantify improvement in performance when a  classification problem is broken down into subproblems in a systematic way,  (b) quantify the advantages of combining different types of classifiers,  create a library of reusable neural network classification models, data  pre-processing, and evaluation tools that use standardized interfaces, and  (d) foster dissemination of classification models and the use of  pre-processing and evaluation tools by making them available to other  researchers through the World-Wide-Web. We will test four hypotheses: (1)  Combinations of different modalities of classifiers perform significantly  better than isolated models. (2) Hierarchical and sequential neural  networks perform better than standard neural networks. (3) Unsupervised  models can decompose a problem for hierarchical or sequential neural  networks better than models that use prior knowledge. (4) It is possible to  build a Classification Tool Kit composed of data pre-processing modules,  classification models, and evaluation modules in which components are  independent, reusable, and interchangeable. | https://reporter.nih.gov/project-details/2385272 | R01 | 1998 | 0.093 |
| The long term goal of our research is to understand the flow of  information from the genome to the phenotype of organisms. In this  proposal, we will attempt to use Bayesian networks and near-optimal  sequence alignments to represent protein secondary structures and motifs.  A Bayesian network describes the likelihood of amino acids at each  position in a motif as well as the dependence of amino acids in one  position on the amino acids at other position. Hence, Bayesian networks  can describe both the conservation of amino acids at single positions and  the conservation of correlations between two positions simultaneously.    Conserved amino acids result from evolutionary selection for a specific  amino acid or type of amino acid at one position in a protein structure.  These positions often have important functional or structural  requirements. Correlated changes between amino acids generally result from  side-chain side-chain interactions between pairs of amino acids in a  protein's structure. The types of correlations we have represented with  Bayesian networks include electrostatic charges, hydrophobicity, hydrogen-  bond donor and acceptor and inversely correlated packing volumes among  others. These Bayesian networks can be used to 1) discover side-chain  side--chain interactions within protei motifs and 2) to search sequence  databases for motifs showing both correlations and conserved amino acids.    Near-optimal alignments between two sequences can display regions that  have been more highly conserved or less highly conserved using the  information contained in only two sequences. The most highly conserved  region correspond to the most highly structured regions and the most  highly variable regions correspond to loops and coils and other  hypervariable regions. We propose to use near-optimal alignments to  display conserved secondary structures of proteins and hypervariable  regions. We will use secondary-structure specific amino acid substitution  matrices to provide specificity.    The goals of this proposal are to 1) build a database of Bayesian networks  that represent protein motifs, 2) test these networks for their ability to  detect motifs using test sets and crossvalidation methods, 3) compare  these networks with other methods for searching protein databases , 4)  build an integrated set of Bayesian networks to predict protein secondary  structure, 5) compare the prediction of protein secondary structure with  existing method 6) build a near-optimal sequence alignment workbench, and  7) predict structured and unstructured regions in proteins from near-  optimal alignments. | https://reporter.nih.gov/project-details/6146063 | R01 | 1999 | 0.093 |
| The goal of this research is to develop an open and extensible software  environment for medical image segmentation. This environment will  contribute to the public and scientific interest in at least three ways:  (1) improved and efficient segmentation of medical images for various  applications, (2) efficient creation of new image segmentation  algorithms, and (3) improved evaluation for medical image segmentation  algorithms.  Image segmentation has many applications in medical imaging; however,  it's use in current clinical practice falls far short of its potential.  Commercially available tools are either designed for very specific  applications, or are general-purpose image processing packages with  little support for image segmentation. The proposed environment will be  devoted to medical image segmentation and will have an extensible and  open architecture. The extensible architecture will allow easy  customization for specific applications and will also allow users to add  their own algorithms. The open architecture will make the software  platform-independent and will allow easy integration with existing  applications, such as databases, analysis packages, or visualization  tools. This environment will also provide tools for evaluation of  medical image segmentation algorithms. To design such a software  environment, we will use the latest innovations in design such a software  environment, we will use latest innovations in software technology, such  as object-oriented design and distributed objects.  PROPOSED COMMERCIAL APPLICATION:  We envision two types of users for this software environment--(1) medical  imaging researchers or medical imaging solution providers, and (2)  clinicians or clinical researchers. The first type of uses will be able  to customize the application completely. The second type of users will  use the software for specific applications. Our building-block approach  to the design of the environment will allow rapid customization for  different applications and for the different types of users. | https://reporter.nih.gov/project-details/2012643 | R43 | 1997 | 0.085 |

**Cluster 13:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The electronic medical record (EMR) holds great allure to both the  medical informatics and health services research communities. In this  project, we propose to enhance the capability of electronic medical  record (EMR) systems by creating and evaluating tools to extract  clinical vocabularies as well as patient data from narrative text  reports. We will apply advanced natural language processing tools from  the CLARIT system to both of the above problems. We contend that fast  and robust automated text processing methods are the only way that the  problems of vocabulary construction and narrative text extraction can  be solved.    We will address the clinical vocabulary problem by utilizing the  thesaurus extraction techniques already present in the CLARIT system.  Using several gigabytes of narrative text, including discharge  summaries, progress notes, radiology reports, and other clinical text,  we plan to:  l. Identify empirically the terminology used in medicine.  2. Compare the coverage of that terminology in several existing large  medical vocabularies: UMLS, SNOMED, and the Medical Entities  Dictionary.  3. Discern the semantic characteristics of that terminology to allow  other structured vocabularies a richer substrate of terms as well as  providing us the opportunity to implement a clinical vocabulary schema  based on the methods of the MedSORT-II Project.  4. Evaluate how well our tools assist the vocabulary building efforts  of ourselves and others.    The narrative extraction problem will be approached differently than  in the past, building on the efforts of previous investigators who  have tackled this problem before but changing the perspective by  focusing on the development of tools specific to researchers and  others with a need to extract data from narrative text. This approach  will be applied in two domains:  l.Consortium-based research in the use of esophogastroduodenoscopy  (EGD).  2.Practice guidelines implementation in blood product transfusion. | https://reporter.nih.gov/project-details/2238284 | U01 | 1995 | 1 |
| Non-small cell lung cancer (NSCLC) is the leading cause of cancer morality  in men and women in the United States, and the overall long-term survial is  less than 15%. Pathologic stage I makes up 25-35% of NSCLC cases and has  a good prognosis. However, cancer relapse and death rate in this subset is  35 to 50% by 5 years. Chemotherapy is beneficial for the treatment of  several localized solid tumors after resection and may prove to be useful  in the treatment of patients with stage I NSCLC. Thr purpose of this  project is to define tissue and serum tumor markers in patients with stage  I NSCLC which predict for early cancer recurrence. Pathologic stage I  NSCLC was chosen for study to eliminate the significant influence of  positive lymph nodes and distant metastases on survival.  Immunohistochemical staining will identify potential tissue tumor markers  and radioimmunoassay (RIA) or enzyme-link immunosorbent assay (ELISA) will  identify potential serum tumor markers.    Specific aim #1 will examine a set of twelve tissue tumor markers in a  retrospective cohort of 275 stage I NSCLC patients. Markers are  categorized by hypothetical method of action: molecular genetic markers  (Kras, erbB-1, erbB-2, rb, p53, bcl-2), markers of metastatic propensity  (angiogenesis factor viii), proliferation markers (K1-67) and markers of  cellular differentiation (Blood group A, H/LeV/LeB, NCAM, CD44). Results  will be used to develop a prediction rule for recurrence in stage I NSCLC  using Cox proportional hazards regression analysis and an artificial neural  network.    Specific aim #2 will examine a set of eight serum tumor markers in a  retrospective cohort of 250 patients with stage I to IV NSCLC. These  markers are categorized as molecular genetic markers (anti-p53), markers of  metastatic propensity (angiogenesis bFGF), somatamedins (growth factor IGF-  1) and markers of cellular differentiation (CEA, CA-125, CA 15-3, CYFRA21-  1, CD44). The purpose of this aim is to identify any correlations between  titers of serum markers and tumor histology, stage or mass. One hundred  patients in this cohort had a second serum collection after tumor  resection. This subgroup of serum will allow analyses of titters before  and after cyto-reduction. Significant correlates with tumor stage and mass  will be evaluated in a prospective cohort of patients with stage I NSCLC.    In specific aim #3, paraffin-embedded and fresh-frozen tumor tissue will be  collected from a prospective cohort of 330 patients with stage I NSCLC to  validate the prediction rule developed in specific aim #1. In these same  patients, serial serum specimens will be collected for a minimum of 2.0  years after resection (specific aim #4). The significant markers  identified in specific aim #2 will be analyzed in this cohort to describe  correlations with tumor recurrence. Tissue and serum markers identified by  the model can be used to select high risk patients for a prospective,  multi-institutional chemotherapy trial for stage I NSCLC. | https://reporter.nih.gov/project-details/2517794 | R29 | 1997 | 1 |
| The long-term objective is to develop computer technology needed to  accomplish the objectives of the Human Genome Project and to apply the  technology to the analysis and management of sequencing data. Currently,  a database search for sequence similarities represents the most direct  computational approach to the analysis of genomic information. However,  the search is becoming ever more forbidding due to the accelerating  growth of sequencing data. The goal of the proposed research is to  further develop and enhance a software tool for speedy classification of  unknown sequences, and make it available to the genome community. The  research will build upon a pilot system designed and developed by the  principal investigator that has shown great promise. The specific aims  are (1) to enhance the tool for speedy identification of PIR  superfamilies and ProSite patterns, (2) to develop a pilot DNA/RNA  classification system, (3) to distribute the tool, and (4) to aid PIR  protein database and RDP ribosomal RNA database organization. In  contrast to other search methods whose search time grows linearly with  the number of entries in the database, the time of the proposed tool  grows with the number of families, which is likely to remain low. The  tool would automate family assignment which is especially important for  managing the influx of new data in a timely manner.    The proposed research applies neural network technology to solving the  database search/organization problem. The major design principles  involve an encoding schema to extract sequence information and a modular  architecture to scale up backpropagation networks. The encoding  algorithm is a hashing function similar to the k-tuple method. A pilot  system has been implemented on a Cray supercomputer to classify electron  transfer proteins and enzymes. The system achieves about 90% accuracy  and 50 times speed of other search methods. The speed may be 1000 times  faster than others in a decade if the database continues to grow at the  current rate. In the proposed research, the sensitivity of the tool  would be improved and a full-scale system would be developed. The  automated software tool would be portable at the source code, user  interface, and hardware levels. The system would be updated in  accordance with database releases, and distributed to the research  community via anonymous ftp. The tool would be used to classify PIR  sequences according to superfamilies and to classify ribosomal RNA  sequences according to phylogenetic relations. | https://reporter.nih.gov/project-details/2445394 | R29 | 1997 | 1 |
| Post-traumatic stress disorder (PTSD) is one of the most disabling  psychopathological conditions affecting the veteran population.  Approximately 15.2% of the men and 8.5% of the women stationed in  Vietnam were found to be suffering from PTSD 15 or more years after  their service. In the Atlanta metropolitan area, some 9000 Vietnam  veterans suffer from complete or partial PTSD. The psychological,  social, occupational and economic consequences of the disorder for  patients and their families are devastating. No therapeutic approach  has proven to be consistently effective in the management of combat-  related PTSD. The present proposal intends to exploit the potential  therapeutic effectiveness of recent advances in computer and display  technology referred to as Virtual Reality. Virtual reality exposure  (VRE) takes place in an immersive, computer-driven environment.  Patients would be exposed to virtual Huey helicopters flying them over  the jungles of Vietnam. They will be encouraged to relive their  traumatic memories, much as in standard exposure therapy, but immersed  in Vietnam stimuli. Ultimate control is possible in the virtual  environment, changing levels of intensity of exposure instantly. The  proposed project aims to develop virtual reality exposure therapy for  Vietnam veterans with PTSD, revise and perfect the treatment,  construct a treatment manual, and gather preliminary evidence of its  efficacy in a small group design. A series of five case studies will  be run to develop and revise the treatment. Following this, Vietnam  veterans (n=40) with current DSM-IV PTSD diagnoses will be randomly  assigned to VRE or a wait-list control. Treatment will be delivered  in nine 60-minute individual sessions conducted over 5 weeks.  Assessments will be conducted at pre-treatment, post-treatment and  follow-ups of 6 and 12 months post-treatment. Assessments will be  conducted by an independent assessor who will be kept blind to the  treatment condition. Objective clinician-rated and self report  measures of PTSD will be incorporated. | https://reporter.nih.gov/project-details/2635521 | R21 | 1998 | 1 |
| DESCRIPTION:(Adapted from the Investigator's Abstract) The goal of this  grant is to determine if and when cancer risks can be estimated by  establishing record-linkages between statewide cancer surveillance systems  and occupational cohorts. More specifically, the aims of this study are to:  (1) determine the feasibility of utilizing statewide cancer surveillance  systems in the evaluation of cancer incidence within occupational cohorts;  (2) compare and contrast the relative merits of standardized incidence  ratios (SIR) with standardized mortality ratios (SMR) as determined from  cancer surveillance systems incidence data and death certificate mortality  data, respectively; and (3) provide recommendations concerning how and when  statewide cancer surveillance systems should be utilized in the evaluation  of occupational cohorts.    SMR and SIR estimates will be calculated and compared for three occupational  cohorts (e.g., Highway Maintenance, 3M, Conwed). SMR analyses have already  been completed for the Highway Maintenance and 3M cohorts; a mortality  update and SMR analysis will be conducted for the Conwed cohort (1988-1995).  Cancer morbidity information, for the SIR analysis, will be determined by  linking the three cohorts with the Minnesota Cancer Surveillance System  (MCSS). Residency status will be required before person-years can be  calculated, however, because inclusion in the MCSS is restricted to  Minnesota residents. Linkages to other data sets will be used to determine  the Minnesota residency status. Sensitivity analyses will be used to  evaluate confounding and follow-up bias. Standardized mortality ratios will  be compared to standardized incidence ratios for the Highway Maintenance,  3M, and Conwed cohorts. Finally, this study will evaluate the utility and  limitations of cancer surveillance systems as a tool for occupational cancer  research; recommendations for its use will be developed. | https://reporter.nih.gov/project-details/6071508 | K01 | 1999 | 1 |

**Cluster 14:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The long-term objective of our research group is to facilitate automatic  or semi-automatic classification and retrieval of natural language texts,  in support of reducing the cost and improving the quality of computerized  medical information. This proposal develops further and applies a novel  approach, the Linear Least Squares Fit (LLSF) mapping, to document  indexing and document retrieval of the MEDLINE database. LLSF mapping is  a statistical method developed by the PI for learning human knowledge  about matching queries, documents, and canonical concepts. The goal is to  improve the quality (recall and precision) of automatic document indexing  and retrieval, which cannot be achieved by surface-based matching without  using human knowledge or thesaurus-based matching dependent on manually  developed synonyms. This project applies LLSF to MEDLINE, the world's  largest and most frequently used on-line database, to evaluate the  effectiveness of this method and to explore the practical potential on  large scale databases. The specific aims and methods are:    l. To collect data needed for the training and evaluation of the LLSF  method. A collaboration with another research institute is planned for  utilizing and refining a large collection of MEDLINE retrieval data. A  sampling of MEDLINE searches at the Mayo Clinic will be employed for  obtaining additional tasks.    2. To develop automatic noise reduction techniques for improving both the  accuracy of the LLSF mapping and the efficiency of the computation. A  multi-step noise reduction in the training process of LLSF will be  investigated, including a statistical term weighting for the removal of  non-informative terms, a truncated singular value decomposition (SVD) for  reducing the noise at the semantic structure level, and the truncation of  insignificant elements in the LLSF solution matrix for noise-reduction at  the level of term-to-concept mapping.    3. To scale-up the training capacity for enabling the LLSF to accommodate  the large size of MEDLINE data. A split-merge approach decomposes a large  training sample into tractable subsets, computes an LLSF mapping function  for each subset, and then merges the lcal mapping functions into a global  one.    4. To improve the computational efficiency by employing algorithms  optimized for sparse matrices and for noise reduction. The potential  solutions include the Block Lanczos truncated SVD algorithm which can  reduce the cubic time complexity of standard SVD (on dense matrices) to a  quadratic complexity, a QR decomposition which solves the LLSF without  SVD, a sparse matrix algorithm which has shown a speed-up in matrix  multiplication and cosine computation by a factor of l to 4 magnitudes,  and parallel computing.    5. To evaluate the effectiveness of LLSF on large MEDLINE document sets  and compare with the performance of alternate indexing/retrieval systems. | https://reporter.nih.gov/project-details/2238093 | R29 | 1995 | 0.58 |
| THIS IS A SHANNON AWARD PROVIDING PARTIAL SUPPORT FOR THE RESEARCH  PROJECTS THAT FALL SHORT OF THE ASSIGNED INSTITUTE'S FUNDING RANGE BUT  ARE IN THE MARGIN OF EXCELLENCE. THE SHANNON AWARD IS INTENDED TO PROVIDE  SUPPORT TO TEST THE FEASIBILITY OF THE APPROACH; DEVELOP FURTHER TESTS  AND REFINE RESEARCH TECHNIQUES; PERFORM SECONDARY ANALYSIS OR AVAILABLE  DATA SETS; OR CONDUCT DISCRETE PROJECTS THAT CAN DEMONSTRATE THE PI'S  RESEARCH CAPABILITIES OR LEND ADDITIONAL WEIGHT TO AN ALREADY MERITORIOUS  APPLICATION. THE ABSTRACT BELOW IS TAKEN FROM THE ORIGINAL DOCUMENT  SUBMITTED BY THE PRINCIPAL INVESTIGATOR.    DESCRIPTION: (adapted from the application abstract) The advent of  large picture archiving and communication systems (PACS) will likely  result in a conversion of clinical radiology to become nearly completely  digital. Improved methods for access to a large collection of on-line  image data can improve medical research and education. Although the  techniques for fast text search are well established, similar tool  development for rapidly searching through years of digitally archived  images is still in its infancy. Development of a fast browsing technology  for retrieving archived images over a network, both local and remote  (international) accesses is proposed.    In this project, three key technologies will be merged: smart query,  hierarchical archive, and photo indexing using embedded zerotree wavelet  transform (EZTWT) code. The first two technologies provides query  constraint and automatic image migration management, and are supported  in other PACS- related projects at UCLA. The third and newest component,  EZTWT, is an encoding method designed for maximum network utility,  providing the receiver highest image quality for a given transmission  time. The implications for high-performance teleradiology under  bandwidth limitations are significant. Diagnostic viewing can start  before transmission of an image in full resolution and quality is  complete. This encoding scheme will be modified to allow variable size  iconic representation of the original, achieving minimum network delay  and fast reconstruction, thus making browsing through a larger selection  of images convenient and manageable.    In addition, the modified EZTWT can achieve high efficiency for images  with clearly separated anatomical and background regions. The space  saving in very high resolution radiographs, such as mammograms, can be  typically 7MB per 4Kx4K image (44%) without loss in anatomical  information. This coding technique will be tested in conjunction with  new pattern recognition techniques to achieve efficient mammogram storage  and telecommunication. | https://reporter.nih.gov/project-details/2329566 | R55 | 1996 | 0.58 |
| The long term goal of this research is to formulate a more comprehensive  model of the development of visual and auditory word recognition. To  achieve this aim, we must investigate the role that different factors  play in the identification of words by children as well as by adults  (e.g., association-strength, graphemic-similarity, phonemic-similarity,  relatedness-proportion, stimulus-quality, and word-frequency). The  specific aim of this proposal is to elucidate the developmental changes  that occur in the processing within and between the grapheme-phoneme,  lexical, and semantic systems. More specifically, we propose to  investigate: (l) whether spreading activation within and between the  lexical and semantic systems is under strategic control; (2) the nature  and time course of the influence of phonemic and semantic information on  the lexical system; and (3) the time course of processing within the  grapheme-phoneme system and the influence of the lexical system on its  output. | https://reporter.nih.gov/project-details/2403049 | F32 | 1997 | 0.58 |
| An Expert-Driven Lung Nodule Detection (E-HLND) System is proposed for  improving diagnostic accuracy and speed for lung cancerous pulmonary  radiology. The research goal is to develop a robust, user-friendly, and  clinically useful system to assist radiologists in the detection and  analysis of lung tumor in an early and treatable stage. The detection and  treatment of lung nodule in the early stage of growth can results in a  better prognosis for survival. The proposed E-HLND system configuration  include the following processing phases: (1) data acquisition of a large  clinical screening chest x-ray films and multiresolution pre-processing  to enhance object-to-background contrast, (2) quick selection of suspect  nodule areas, (3) features space determination and neural classification  of cancerous nodules as well as false positives, and (5) knowledge-based  registration and fusion processing to integrate follow-up, patient  history, radiologist's expertise, and other diagnosis reports. This  proposal focuses on (l) improving system's sensitivity and specificity  with neural network, image registration, information fusion technologies,  and hardware design; and (2) validating system's performance with a  "simulated" clinical trials based on a large clinical x-ray film database  (Chinese Yunnan Tin Corp. Bio-Marker Specimen bank -YTC database). This  project will explore artificial neural network and computer vision  technologies in diagnostic radiology and provide a basis for other cancer  research in diagnostic radiology. This R&D effort is not only consistent  with but its success will provide good tool in the NCI launched large-  scale study "Prostate, Lung, Colorectal, and Ovarian Cancer Screen (PLCO)  Trial".    PROPOSED COMMERCIAL APPLICATIONS  An expert-driven lung nodule detection system, which serves as a "second  reader" to assist pulmonary radiologists in detecting lung nodules, will  be of great clinical and commercial value. The system can increase  radiologists' sensitivity and specificity in the detection of early lung  cancer on screening chest radiographs. Early and accurate detection of  an early stage tumor will ensure patients get the best treatment  available. The proposed R&D work will enhance current patient care  system, reduce the work load of radiologists, and improve the cancer  diagnostic procedure in diagnostic radiology. | https://reporter.nih.gov/project-details/2542475 | R44 | 1998 | 0.51 |
| Atrial arrhythmias are extremely common and are often associated with  significant morbidity and mortality. The ability to develop and  prescribe new treatments for atrial arrhythmias depends on the  availability of better diagnostic tools that are simple, inexpensive,  and effective enough to support their use for screening large numbers  of patients.    During Phase I, we developed signal processing algorithms that permit  noninvasive assessment of patients suffering from atrial fibrillation.  T algorithms were tested on patient data and found to assist clinicians  in predicting patient response to cardioversion in an attempt to  terminate the arrhythmia and restore sinus rhythm.    During Phase II, we will investigate alternate algorithms for  identification and classification of atrial arrhythmias. Our goal is  development of a tool that will allow the clinician to rapidly identify  specific arrhythmias and measure the progress of individual patients  through different courses of therapy.    PROPOSED COMMERCIAL APPLICATION:  The technology developed under this program could be used in diagnostic  devices like electrocardiogram monitors. | https://reporter.nih.gov/project-details/2777213 | R44 | 1999 | 0.51 |

**Cluster 15:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| This grant proposal is focused on developing new mathematical and  statistical models to describe biological systems. Models to represent,  help to understand, predict future behavior, and control biological  systems are becoming more and more important and of widespread use in  different fields related to biology and health care. Complex mathematical  models are needed to model the complicated interactions between the  physiological functions of biological systems, and to model the effect of  interventions (e.g. therapy) on these functions. The specific aims of this  grant focus on three areas of research. 1. Develop and investigate  statistical models for biological population data. Biological data are  always collected from some population of different individuals, and are  often highly variable. This is mostly due to variability of physiological  functions between individuals, and to measurement error. Statistical  models are needed to deal with the complex structure of population data. I  will (I) introduce a general methodology based on the use of sophisticated  heteroscedastic statistical models, which does not explicitly formulate a  model for interindividual variability but promises to be fast, efficient  and unbiased; and (ii) investigate the performance of existing population  models using realistic simulations including model misspecification. 2.  Develop semi-mechanistic compartmental models. I focus on three main  problems: (i) the development and investigation a new general class of  compartmental pharmacokinetics"'pharmacodynamic (PK/PD) models, (ii) the  development of semi-mechanistic black-box compartmental models to deal  with non-linear PK systems, (iii) the development of the technology to  apply well established semi-mechanistic linear black-box models to the  purpose of PK control. 3. Develop new multivariate dynamic models. The  main problem addressed is how to represent a system where multiple inputs  (drugs) and multiple interrelated responses are measured. I propose  different classes of models to do so based on spline networks and  eventually neural networks. The proposed models can incorporate a  compartmental sub-structure to easily deal with kinetics. Continuous and  discrete time versions of the models are considered. The statistical and  mathematical models introduced in the grant have widespread application to  a variety of biological fields. However specific areas, directly linked to  health care issues, are selected for active research and application of  the proposed models. These areas correspond to experimental situations  where the models proposed in the grant are particularly needed (nonlinear  and multivariate dynamic), and represent continuations of already  established collaborations with leading scientists. They include: computer  control of ultra-short acting anaesthetic drugs administration,  pharmacokinetics/pharmacodynamic of short-acting anesthetics,  pharmacodynamic of nicotine and nicotine tolerance development, adenosine  kinetics and metabolism and their relationship to adenosine  pharmacodynamic effects, modeling of cardiovascular drugs effects on  pharmacy dynamic responses (heart rate, blood pressure, and breathing  variability) sampled at high rates. | https://reporter.nih.gov/project-details/2189554 | R29 | 1995 | 0 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**Cluster 16:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
|  | https://reporter.nih.gov/project-details/2040081 | R44 | 1997 | 0.89 |
| Efforts to apply computer methods to assess and improve the quality of  care in the hospital have been stymied by limited access to clinical  data. Free-text data have detailed clinical descriptions of patients  that would be useful in computer altering systems and computer reminder  systems. However, free-text data cannot be interpreted by most clinical  computer systems. In this proposal, we describe research specifically  aimed at making free-text data accessible to computer-based applications  for assessing and improving the quality of care. In particular the  research plan focuses on the development of technologies that would allow  free-text data to be used in clinical alert systems for critical test  results; in reminder systems to encourage adherence to practice  guidelines; and in data collection systems for severity of illness models  applied in the assessment of risk adjusted outcomes. The approach  described in the research plan emphasizes the development of statistical  and probabilistic methods for interpretation of data derived from medical  language processing systems. We will test the methods developed for  language processing and interpretation developed under this proposal in  3 area: 1) the identification of concepts related to severity from the  MedisGroups and the Computerized Severity Index models of patient  severity of illness; 2) the identification of chest x ray reports and  mammography reports with potentially malignant findings that require  radiological follow-up; 3) and the automatic assessment of  appropriateness of coronary artery bypass grafting (CABG) surgery from  free-text descriptions of patients based on the application of a clinical  practice guideline for CABG surgery. | https://reporter.nih.gov/project-details/2909061 | R29 | 1998 | 0.89 |
| DESCRIPTION: Computing with biochemical reactions is increasingly important  in studying genomes, assessing toxicity, and developing therapeutics. There  are several important information sources, but their data are rudimentary  and often inaccurate. Incorporation of biochemical information into  databases is extremely slow compared to that of sequence and structural  information, and will lag further as large-scale surveys of gene expression  and other reactions accelerate over the next few years. Mechanisms for  review exist, but are manual, paper-dependent, and can be delayed for a year  or more.    As curators and coordinators of biochemical information sources, the  applicants share a number of problems in the collection and review of  information. Moreover, they are mutually dependent for the means to do so:  compound information is critical in checking reaction data, reaction  information is needed to spot errors in compound information, and the  automatic verification algorithms for either are closely related and need  both. The applicants, therefore, propose to build a curatorial exchange for  the deposit and review of biochemical information by the scientific  community. The applicants' goal is to demonstrate a system that will  encourage the mandating of deposit while ensuring that the information is of  the highest quality.    The role of the exchange is to receive deposits, check and classify their  biochemical information automatically, forward them to panels of human  reviewers for vetting, and publish the information by release to the  participating data sources--all over the World-Wide Web. It will track the  origin and status of deposits and reviews, serve computations for the  relevant pattern matching and simulation, and maintain an archival copy of  data. The databases remain independent, and separately provide additional  information. Algorithm development and testing depends on an adequate  information infrastructure, so the applicants will complete a basic data set  of compounds and reactions. They will use this experience to develop a more  comprehensive domain model that better captures modern biochemistry, and  implement it for deposit and review. Since the basic data and algorithms  will be valuable to the community at large, they plan to serve these to the  World-Wide Web. the exchange and its underlying data form the infrastructure  necessary for sustainable, cost-effective development of biochemical  informatics resources for biomedical research. | https://reporter.nih.gov/project-details/2771106 | R01 | 1998 | 0.89 |
|  |  |  |  |  |
|  |  |  |  |  |

**Cluster 17:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| Changes in soft tissue elasticity are usually related to pathological  processes. Because of this, palpation is still widely used for  diagnosis. Its efficacy, however, is limited to abnormalities located  relatively close to the skin surface. The goal of quantitative  elasticity imaging is to develop surrogate, remote palpation, thus  expanding its range to include deep lying lesions. The elastic  properties of any continuous medium such as tissue can be assessed  through precise measurement of mechanical deformations throughout that  medium induced by forces applied at the surface. Using modern medical  imaging devices to precisely measure internal motion, it should be  possible to estimate and even image elastic properties of internal  organs. In competition with other imaging modalities, ultrasound has two  major advantages for elasticity imaging; it is inherently real-time and  speckle artifacts limiting the quality of conventional images provide  excellent markers for accurate tracking of tissue motion. Elasticity can  be imaged, therefore, by measuring motion with an ultrasound speckle  tracking algorithm, followed by reconstruction of the elasticity  distribution. Although some other imaging systems, particularly real-  time ultrasound, must be used to monitor tissue motion, elasticity  imaging represents a fundamentally new diagnostic modality. To  investigate quantitative elasticity imaging for medical diagnosis, a  research plan addressing the important clinical problem of renal  inflammation and scarring has been formulated. Preliminary data support  the hypothesis that kidney elasticity changes with renal damage and  concomitant scarring before renal problems are detectable by traditional  diagnostic techniques such as laboratory measurements of renal function.  Therefore, quantitative elasticity imaging may be valuable in detecting  and quantifying scar for conditions such as kidney transplant rejection  where rejection is difficult to quantify from functional measurements  alone. Based on the results of these studies, it is the long range goal  of this research program to develop a sensitive diagnostic technique  based on quantitative elasticity imaging permitting surrogate palpation  of deep lying lesions. | https://reporter.nih.gov/project-details/2146817 | R01 | 1995 | 0.019 |
| DESCRIPTION: Over the past decade a variety of alternative computer  based modeling techniques have been introduced which show  promise for the construction of clinical decision aids. These  techniques include statistical regression approaches such as  generalized additive modeling, classification tree induction such as  ID3 or CART, and multi-layer neural networks. Logistic regression  models (LR) are currently central to most probabilistic predictive  clinical decision aids and are fundamental to comparative analyses of  medical care based risk adjusted events. These newer techniques  have been applied on a larger scale in the last few years. They  appear to have unique advantages in selected circumstances. The  successful use of these methods, however, depends on understanding  their accuracy, performance, and model transportability.    A formal assessment of these new techniques with four specific aims  is proposed: (1) to assess and compare the performance of  different models to determine the factors which affect  performance; (2) to develop automated computer based procedures  for exploratory model development for each method; (3) to develop  hybrid models incorporating the strengths of each of the existing  techniques, and (4) to determine the situations that restrict the  transportability of these models.    These specific aims will be achieved in a three stage project. In  the first stage four approaches will be pursued: (1) the  mathematical properties of the different computational algorithms  for the modeling techniques will be studied; (2) automated  modeling procedures will be developed and utilized; (3) the factors  that affect performance for each modeling technique will be  explored and(4) new hybrid techniques will be developed and assessed.  In the second stage the methods developed in the first stage will  be used to create and test models that predict cardiovascular  events on data from 15,000 patients in a prospective clinical  trial. In the third stage the factors that affect the  generalizability and transportability of models to new datasets will  be explored by repeated sampling and model construction on  different subsets of the cardiovascular database including  separating the database into subsets from each of ten different  hospitals.    This work will broaden the understanding of these important  modeling techniques and their potential contributions for  clinical decision making, health policy research, and medical  informatics. New modeling techniques might be developed which  incorporate elements from different techniques. | https://reporter.nih.gov/project-details/2237926 | R01 | 1995 | 0.019 |
| The goal of this proposal is the development of Computer-Aided  Instrument Design (CAID) tools and resources for the creation of  embedded biomedical instruments. The Phase I work has demonstrated  the conceptual feasibility of our approach through the development  of a programmable, portable, hardware platform and a reusable set  of software libraries. These tools and libraries were used to create  a prototype of a Cardiac Monitor instrument. The Phase II work is  aimed at expanding and improving the existing rudimentary CAID tools  and resources to create a complete, turnkey environment for the  design of many different physiological instruments. The Phase II  work will additionally involve creating and testing two  representative instruments: one will be an advanced ECG research  device suggested and defined by our collaborators at JHU (Johns  Hopkins University); the other will be an Airway Monitor, suggested  and defined by our collaborators at EVMS (Eastern Virginia Medical  School).    PROPOSED COMMERCIAL APPLICATIONS The potential commercial  opportunities of our proposed technology are compelling. In addition  to our selected Phase II experimental projects, various potential  collaborators suggested applications including: an advanced Holter  monitor with real-time detection of specific arrhythmias, fetal  heart monitors (using acoustic-based sensors of different  technologies), vital signs acquisition and logging,  ambulatory/portable EEG devices for sleep studies, seizure, and  brain injury detection, and many others. | https://reporter.nih.gov/project-details/2771527 | R44 | 1998 | 0.019 |
| During Phase I, Flint Hills Scientific developed an algorithm for real  time quantitative seizure detection which performs with sensitivity and  specificity equal to expert visual analysis. Of even greater value is the  capability of this algorithm to predict seizure onset by 13.6 seconds  (mean), in its generic mode. Preliminary studies indicate that with  automated individualized adaptation, prediction time can be increased to  180 seconds or longer. To the best of our knowledge no other system in  existence has achieved this level of success. These results lay the ground  for the fulfillment of "seizure prediction, early recognition and blockage  of seizures," the number one AES research priority.    The main goal of Phase II will be to advance, further refine, and validate  this technology for implementation into a portable or implantable device  with diagnostic, warning, and therapeutic capabilities. We are confident  that this technology, by decreasing or eliminating unpredictability, will  minimize the potentially devastating effect of seizures on quality of life  while decreasing morbidity, the cost of health care, and the reliance on  the welfare system. These unique advantages will ensure widespread  acceptance of this technology by those directly and indirectly affected by  epilepsy and by the health care system.    PROPOSED COMMERCIAL APPLICATION:  1. Software package for real time seizure prediction, detection,  localization, imaging, and quantitative analysis. 2. Software package for  automated, selective noise reduction 3. Portable device for the automated  early warning of impending seizures. 4. Portable or implantable devices  for automated early therapeutic intervention. | https://reporter.nih.gov/project-details/2771953 | R44 | 1998 | 0.019 |
| DESCRIPTION: (Adapted from investigator's abstract) This project will  examine new methodology for making inference about the regression parameters  in the presence of missing covariate data for two commonly used classes of  regression models. In particular, we examine the class of generalized  linear models for general types of response data and the Cox model for  survival data. The methodology addresses problems occurring frequently in  clinical investigations for chronic disease, including cancer and AIDS. The  specific objectives of the project are to: 1) Develop and study classical  and Bayesian methods of inference for the class of generalized linear models  (GLM's) in the presence of missing covariate data. In particular, we will  i) examine methods for estimating the regression parameters when the missing  covariates are either categorical or continuous and the missing data  mechanism is ignorable. Also, parametric models for the covariate  distribution will be examined. The methods of estimation will focus on the  Monte Carlo version of the EM algorithm (Wei and Tanner, 1990) and other  related iterative algorithms. The Gibbs sampler (Gelfand and Smith, 1990)  along with the adaptive rejection algorithm of Gilks and Wild (1992) will be  used to sample from the conditional distribution of the missing covariates  given the observed data. ii) examine estimating the regression parameters  when the missing covariates are either categorical or continuous and the  missing data mechanism is nonignorable. Models for the missing data  mechanism will be studied. iii) develop and study Bayesian methods of  inference in the presence of missing covariate data when the missing  covariates are either categorical or continuous and the missing data  mechanism is ignorable. Parametric prior distributions for the regression  coefficients are proposed. Properties of the posterior distributions of the  regression coefficients will be studied. The methodology will be  implemented using Markov Chain Monte Carlo methods similar to those of  Tanner and Wong (1987). iv) investigate Bayesian methods when the  covariates are either categorical or continuous and the missing data  mechanism is nonignorable. Multinomial models for the missing data  mechanism will be studied. Dirichlet prior distributions for the  multinomial parameters will be investigated.    2) Develop and study classical and Bayesian methods of inference for the Cox  model for survival outcomes in the presence of missing covariates.  Specifically, we will i) develop and study estimation methods for the Cox  model for survival outcomes in the presence of missing covariates. Methods  for estimating the regression parameters when the missing covariates are  either categorical or continuous will be studied. The methods of estimation  will focus on an EM type algorithm similar to that of Wei and Tanner (1990).  ii) study estimation of the regression parameters when the missing  covariates are either categorical or continuous and the missing data  mechanism is nonignorable. Models for the missing data mechanism will be  studied. Bayesian methods similar to those of 1-iii) and iv) will be  investigated. Computational techniques using the Monte Carlo methods  described in 1-iii) will be implemented. | https://reporter.nih.gov/project-details/2769933 | R01 | 1998 | 0.019 |

**Cluster 18:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
|  | https://reporter.nih.gov/project-details/2303860 | N44 | 1993 | 1 |
| A significant fraction of all courses of external-beam radiotherapy have  unplanned interruptions, i.e. scheduled treatments are missed. Because of  the kinetics of cellular damage repair and tumor repopulation, designing  extra or augmented treatments to compensate for the missed fractions is  non-trivial. Since some individualized compensation regimen must be  chosen, it should be rationalized using state-of-the-art radiobiology. In  Phase I we wrote prototype software (RIC, the Radiotherapy Interruption  Compensator) which produces a range of practical options for the  radiotherapist to compensate for unplanned interruptions during an  extended radiotherapy regimen. It calculates what doses are needed to  produce equal tumor control (or, alternately, equal late normal-tissue  complications) when a proposed interruption-compensated regimen is  substituted for the originally planned protocol. The program is designed  as an "expert system", making available to any practicing radiotherapist  the expertise of leading practitioners, in a personal-computer based,  user-friendly package.    Having demonstrated feasibility, we propose, in this Phase II application,  to continue the research effort initiated in Phase I. The goal is to  produce, at the end of the Phase II period, a marketable product which is  a) scientifically state-of-the-art, b) designed and written with  appropriate quality control, c) robust to use, d) credible to practicing  clinicians.    PROPOSED COMMERCIAL APPLICATION: Potential market consists of more than  2,500 Radiation Oncology facilities worldwide. Potentially about 1-4  copies of the software purchased by each facility, at about $900 per copy. | https://reporter.nih.gov/project-details/2106040 | R44 | 1996 | 1 |
| This acquisition will provide technical assistance to continue the  development and enhancement of GRATEFUL MED as well as provide technical  software development support for the analysis, design, implementation,  integration, documentation, and maintenance of related system components.  These include: computer aided learning, intermachine implementation,  artificial intelligence, new telecommunication interfaces, e.g., FTS  2000, Graphical User Interfaces, network integration and  interoperability, and document delivery systems. | https://reporter.nih.gov/project-details/2394866 | N01 | 1997 | 1 |
| We propose to extend the successful work we have achieved with  statistically based indexing and retrieval systems, by incorporating  semantic structures which accommodate the modifying attributes of clinical  conCepts. Patient data is rarely limited to a single axis of meaning or  detail, and retrieval for application in quality improvement, decision  support, or epidemiologic research, demands Consistent information  struCture. This proposal will invoke the knowledge and tool suites of the  UMLS Specialist Lexicon, the SGML markup and recognition capabilities of  the TextMachine application, extensions to our locally developed CliniCal  Query Language, and layer these enhancements upon our core techniques for  statistically based indexing and retrieval of patient data. We commit  these activities to remain compliant with emerging standards for medical  concept representation arising from the Canon efforts and the  standardization processes at ANSI-HISPP, CEN TC251 and the CPRI  initiatives. | https://reporter.nih.gov/project-details/2635406 | R01 | 1998 | 1 |
| The transport system of the corneal endothelium maintains the cornea at  the low level of hydration required for transparency and good vision. If  this transport system does not function properly then permanent corneal  edema, loss of transparency, and eventual blindness may occur.    Our long term objective is to understand how this transport system  functions under normal conditions and how it changes in aged, injured, and  diseased corneas.    Ion channel proteins in the cell membrane are an essential component of  this transport system. We will study the dynamics of how these proteins  function as "molecular machines".    The specific aims for this project period are:    1) to perform patch clamp experiments to measure and compare the  properties of ion channels from freshly excised and cultured corneal  endothelial cells,    2) to develop and apply new mathematical methods, including powerful new  methods of fractals and nonlinear dynamics (chaos) to analyze this data,  and    3) to determine the properties of different types of dynamical models to  understand the molecular mechanisms responsible for the channel  properties. | https://reporter.nih.gov/project-details/2833074 | R01 | 1998 | 1 |

**Cluster 19:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The amino acid sequence of a protein uniquely determines its tertiary  structure. Deciphering this relationship, the protein folding problem, has  become increasingly important to molecular biologists. DNA sequencing has  become routine, but structural experiments remain very difficult.  Computational strategies are needed to help address this problem.    This proposal describes a strategy to identify the location of  alpha-helices and beta-strands throughout the sequence. A rationale is  offered for employing neural networks and pattern based algorithms to  address the secondary structure prediction problem. Once secondary  structure is located, computational methods exist for generating plausible  tertiary structures. However, these combinatorial strategies give rise to  a large number of alternative structures which are difficult to distinguish  from the correct fold. Simplified potential functions are proposed as a  method for overcoming this structure evaluation problem. The properties of  a non-lattice based simplified representation of a polypeptide chain will  be explored to aid in the construction of an appropriate simplified  potential function. Collaborative ventures are planned to experimentally  test the merits of existing algorithms for predicting protein structure.    In collaboration with Dr. Bunn at Harvard, the relationship of the  erythropoietin sequence to its structure and function will be explored. In  collaboration with Dr. Wang at UCSF, the merits of a proposed structure of  hypoxanthine guanine phosphoribosyl transferase will be studied using site  directed mutagenesis. An exploration of the possibility of grafting the  active site of one enzyme onto the structural scaffold provided by another  protein will be studied in collaboration with Dr. Craik at UCSF and Dr.  Wells at Genentech. | https://reporter.nih.gov/project-details/2180085 | R01 | 1995 | 1 |
| The goal of this project is to refine and evaluate techniques that  automatically construct, from clinical databases, Bayesian belief networks  that can be used as diagnostic and prognostic aids. The amount of  clinical information stored in databases has increased markedly in the  last two decades, and it seems likely that this trend will continue.  Belief networks are able to represent the probabilistic dependencies among  clinical variables in a relatively general manner. Researchers have  developed algorithms for performing probabilistic inference using belief  networks, and they have applied these algorithms to perform medical  diagnosis and prognosis. Although advances have been made in developing  the theory and application of belief networks, the manual construction of  these networks often remains a difficult, time-consuming task. The  automated generation of belief networks from high-quality databases may  facilitate significantly the construction of diagnostic and prognostic  systems, which can serve as clinical decision aids, after their accuracy  and usefulness are validated.    The long-range goal of this research is to advance our understanding and  development of probabilistic systems that can serve as useful diagnostic  and prognostic tools for physicians. Such systems can serve as one method  for disseminating the clinical knowledge captured in high-quality  databases, such as those developed from PORT studies. Within this  context, the specific aims of the current, proposed research project are  to:    \* refine and extend current methods for automatically constructing belief  networks from large databases;    \* test the diagnostic and prognostic accuracy of systems that are based on  belief networks constructed automatically from high quality databases,  compared to several standard statistical techniques;    \* test whether a combination of automated and expert-based methods for  constructing belief networks will yield diagnostic and prognostic systems  that are more accurate than systems that are based on belief networks that  are constructed automatically.    These three aims will be pursued using large, high-quality clinical-  research databases at the University of Pittsburgh that contain  information on patients with syncope and patients in a PORT study with  community-acquired-pneumonia. | https://reporter.nih.gov/project-details/2460259 | R29 | 1997 | 1 |
|  | https://reporter.nih.gov/project-details/2423759 | R44 | 1997 | 1 |
| The growing use of DNA sequence data in research, databases, diagnostic  and therapeutic biotechnology, and even litigation dramatically  increases the need to improve the quality of data being used. This  proposal addresses the problem of assembling a large set of sequenced  DNA fragments into a finished consensus. In order for a sequencing  project to produce high quality finished sequence data, the assembly of  sequence fragments must be correct and accurate both in its large scale  structure and in the fine scale detail of the alignment of individual  base calls. We propose to investigate new algorithms for consensus  estimation and assembly of DNA sequence fragments. Recent novel word-  based approaches to consensus estimation offer promise as a method for  de novo assembly and for exploring alternative assemblies on the large  scale. This will be especially important when sequences contain large  exact or approximate repeats. We propose to develop several main  enhancements to these algorithms. In particular, we will develop a  global optimization algorithm for determining consensus sequences,  replacing current locally optimizing methods. Also, we propose to  develop algorithms allowing alternative alignments in regions of  ambiguity. This approach will allow us to assess alignment accuracy at  both the large and fine scale level.    PROPOSED COMMERCIAL APPLICATION  Accurate assemblies are at the heart of many sequencing projects central  to biopharmaceutical, agricultural, and basic research as well as to the  Human Genome Project. The proposed advances will provide the potential  for simultaneously increasing reliability and automation in a  bioinformatics software market totaling about 100 million dollars per  year. | https://reporter.nih.gov/project-details/2536784 | R43 | 1998 | 1 |
| Efforts to apply computer methods to assess and improve the quality of  care in the hospital have been stymied by limited access to clinical  data. Free-text data have detailed clinical descriptions of patients  that would be useful in computer altering systems and computer reminder  systems. However, free-text data cannot be interpreted by most clinical  computer systems. In this proposal, we describe research specifically  aimed at making free-text data accessible to computer-based applications  for assessing and improving the quality of care. In particular the  research plan focuses on the development of technologies that would allow  free-text data to be used in clinical alert systems for critical test  results; in reminder systems to encourage adherence to practice  guidelines; and in data collection systems for severity of illness models  applied in the assessment of risk adjusted outcomes. The approach  described in the research plan emphasizes the development of statistical  and probabilistic methods for interpretation of data derived from medical  language processing systems. We will test the methods developed for  language processing and interpretation developed under this proposal in  3 area: 1) the identification of concepts related to severity from the  MedisGroups and the Computerized Severity Index models of patient  severity of illness; 2) the identification of chest x ray reports and  mammography reports with potentially malignant findings that require  radiological follow-up; 3) and the automatic assessment of  appropriateness of coronary artery bypass grafting (CABG) surgery from  free-text descriptions of patients based on the application of a clinical  practice guideline for CABG surgery. | https://reporter.nih.gov/project-details/2714212 | R29 | 1998 | 1 |

**Cluster 20:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| New algorithms and computer software tools will be developed to aid in  identifying the function of newly-generated sequences. This work will  have important practical applications for human and model organism genome  sequencing projects. Significant insights into the potential function  of newly-generated sequences of unknown biological function (e.g.,  anonymous cDNAs), can be obtained if similarity to sequences of known  function can be detected. Current sequence database search programs can  fail to detect similarity between distantly related sequences incases  where functional domains contain a few key residues that are dispersed  along the primary sequence (e.g., "zinc-finger" DNA binding domains).  This is because, in the generation of alignment scores, mismatches at  non-conserved residues can easily outweigh matches at the few key sites.  To overcome this problem, we propose to develop new pattern construction  and search methodologies that identify and utilize only conserved  residues and domains in sequence similarity searches. First, techniques  to identify conserved regions within protein sequences will be used to  construct a new type of sequence database in which only the conserved  regions are represented in each sequence. This database should  significantly improve the ability to detect distantly related sequences  by reducing the number spurious, but statistically significant, matches  to unrelated sequences during a database search. Second, methods will  be developed to exploit information on 1)sequence family relationship and  2) the positions of conserved domains within related sequences in  sequence database searches. These new tools will aid in distinguishing  weak matches for distantly related sequences from the alignments of  unrelated but statistically significant matches in database searches.  Third, new pattern libraries will be constructed from sequence and  sequence similarity data available in the Entrez: Sequences database,  produced by the National Center for Biotechnology Information (NCBI).  This will allow functional information in the covering pattern databases  to be directly cross-referenced to sequence and sequence annotation  information in Entrez database, providing value-added benefits for both  databases. Fourth, the high-speed database search tool BLAST will be  adapted for pattern database searches. This will provide a fast and  sensitive search tool for identifying the function of newly-generated  sequences. Fifth, the use of concave gap penalties and suboptimal  alignments will be incorporated into our Pattern-Induced Multi-sequence  Alignment (PIMA) algorithm. These new extensions will significantly  enhance the quality of the patterns and multiple sequence alignments  generated by PIMA. These new analysis tools should prove invaluable to  genome scientists and molecular biologists as they isolate genes and  proteins of unknown biological function. | https://reporter.nih.gov/project-details/2209206 | R01 | 1995 | 1 |
| Synthesizing information from multiple real-time data sources to  generate an intelligent and coherent assessment of a patient's underlying  hemodynamic condition is of vital importance for the development of  intelligent medical monitors. Dr. Dean Sittig is currently directing the  development of a prototype intelligent cardiovascular monitor (ICM). The  proposed project builds on this work and includes: 1) identifying  medically meaningful trends and artifacts from multiple real-time data  sources, 2) generating patient-specific "smart" alarms, and 3) testing  and refinement of algorithms designed to recognize several basic  hemodynamic abnormalities (e.g., hypovolemia or cardiac tamponade, etc.).  This proposed research project will focus particularly on the intelligent  synthesis of information derived from multiple real-time data sources  (i.e., the trend detection, artifact recognition, and physiological state  determination algorithms) to generate a more systematic assessment of a  patient's hemodynamic status.     Finally, we will test a prototype of the ICM using data from the  intensive care unit and the operating room. Trends and artifacts  detected, alarms generated, and hemodynamic abnormalities recognized will  be compared with a "gold standard" annotated medical record generated by  the clinical staff, to help assess the utility of the monitor in terms of  the goals listed above. | https://reporter.nih.gov/project-details/2032337 | R29 | 1996 | 1 |
| The Family Planning Council of Southeastern Pennsylvania, in collaboration  with the Cancer Prevention Research Center of the University of Rhode  Island, is proposing an innovative study designed to address the risk  behaviors associated with a recent alarming increase in the incidence of  cervical neoplasias among young women. The proposed study is designed to  develop, implement and evaluate interventions that increase consistent  condom use and decrease cigarette smoking among 1,800 low income female  youth aged 14-17 years who obtain family planning services at four diverse  federally-funded family planning clinics. Smoking and unprotected sexual  intercourse have been found to be independently associated with increased  risks of cervical cancer in this population The proposed intervention is  based on the Transtheoretical or Stages of Change model combined with one  of the most promising modalities for reaching youth, an interactive  computer based expert system whose efficacy will be evaluated alone and in  combination with an adaptation of the anticipatory counseling model. Data  will be collected at four points during a nine-month intervention period  and at six-month intervals for 18 months post.intervention to assess  effects over time. In addition, the intervention is aimed at increasing  utilization of comprehensive, gynecologic health care including routine  Pap smear screening, follow.up colposcopic examination and treatment of  cervical dysplasia, the precursor of cervical cancers, when indicated.  The proposed study represents the combined expertise of family planning  researchers and service providers with considerable experience working  with economically disadvantaged females and behavioral scientists with  extensive research expertise in high risk behavior change. | https://reporter.nih.gov/project-details/2467284 | R01 | 1997 | 1 |
| An efficient package for augmenting the communication of individuals  with speech and motor impairments, based on the hand-held computers  known as personal digital assistants (PDA s) is proposed. A PDA-based  communicator would provide an inexpensive and truly portable alternative  to the relatively inefficient and costly commercial systems currently  available. The proposed configuration will employ a pared-down version  of the Integrated Modifiable Package of Augmentative Communication  Technologies (IMPACT), a sophisticated PC-based program that  incorporates a variety of established and proprietary techniques for  enhancing communication. IMPACT employs prediction engines that exploit  the redundancy of written language, embedded within a reconfigurable  graphical interface structure which makes possible several innovative  accelerative communication techniques. After investigating the  available platforms and selecting the most appropriate, a reduced  version of IMPACT will be defined, ported to the PDA, and evaluated.  The portable implementation will require minor sacrifices in  functionality, due to the simplified PDA microprocessor and reduce  memory, but the advanced augmentative techniques of IMPACT will still  yield an effective communication device. This will formally established  thorough a thorough functional evaluation regime, followed by empirical  usability trials. To expand the basic test production capabilities of  the communicator, a speech synthesizer peripheral will be integrated  into the design.    PROPOSED COMMERCIAL APPLICATION  The PDA communicator is targeted at the millions of mobile individuals  with disabilities that prevent both speech and writing, providing  augmented communication in an inexpensive and convenient package.  Secondary applications include portable speech production for persons  with only speech and/or hearing impairments, predictive keyboard  emulation for entry of text and data into PDA s, and development of  communication skills for those with learning disabilities. | https://reporter.nih.gov/project-details/2645359 | R43 | 1998 | 1 |
| The main goal of this research is to develop an open and extensible  software environment for medical image segmentation. Image segmentation  has many applications in medical imaging, including tumor localization,  and radiation therapy planning; however its current use falls far short  of its potential. One reason for this is the lack of support for image  segmentation in commercially available general-purpose software tools.  Our proposed research will overcome this limitation by providing a  software environment will include a wide variety of state-of-the-art  segmentation algorithms, and it will be extensible so that users can add  new algorithms easily. Our software environment will further segmentation  algorithm research by providing the basic infrastructure for image  segmentation algorithm development, such as tools for image display,  region-of-interest selection, manual segmentation, and algorithm  evaluation, as well as the data structures for segmentation. In Phase  I of this research, we have developed a prototype of such a software  environment using recent advances in software engineering. In Phase II  we will further develop the prototype to produce a full-featured software  system. We will design and add new segmentation algorithms and expand  the algorithm evaluation capabilities. We will also apply the software  to three important medical problems where reliable image segmentation is  the key requirement.    PROPOSED COMMERCIAL APPLICATIONS:  We envision several families of products from this research: (l) Software  applications for solving specific medical problems, targeted to clinical  specialists. (2) General-purpose segmentation program for clinical  generalists, to be combined with a develop infrastructure kit for use by  software vendors and algorithm researchers. (3) Software components for  use by software vendors to add segmentation capabilities to their image  processing software. (4) Add in Software Modules | https://reporter.nih.gov/project-details/2896010 | R44 | 1999 | 1 |

**Cluster 21:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
|  | https://reporter.nih.gov/project-details/2272263 | R43 | 1995 | 1 |
| The Family Planning Council of Southeastern Pennsylvania, in collaboration  with the Cancer Prevention Research Center of the University of Rhode  Island, is proposing an innovative study designed to address the risk  behaviors associated with a recent alarming increase in the incidence of  cervical neoplasias among young women. The proposed study is designed to  develop, implement and evaluate interventions that increase consistent  condom use and decrease cigarette smoking among 1,800 low income female  youth aged 14-17 years who obtain family planning services at four diverse  federally-funded family planning clinics. Smoking and unprotected sexual  intercourse have been found to be independently associated with increased  risks of cervical cancer in this population The proposed intervention is  based on the Transtheoretical or Stages of Change model combined with one  of the most promising modalities for reaching youth, an interactive  computer based expert system whose efficacy will be evaluated alone and in  combination with an adaptation of the anticipatory counseling model. Data  will be collected at four points during a nine-month intervention period  and at six-month intervals for 18 months post.intervention to assess  effects over time. In addition, the intervention is aimed at increasing  utilization of comprehensive, gynecologic health care including routine  Pap smear screening, follow.up colposcopic examination and treatment of  cervical dysplasia, the precursor of cervical cancers, when indicated.  The proposed study represents the combined expertise of family planning  researchers and service providers with considerable experience working  with economically disadvantaged females and behavioral scientists with  extensive research expertise in high risk behavior change. | https://reporter.nih.gov/project-details/2390828 | R01 | 1997 | 1 |
| This proposal aims to develop an improved understanding of the mechanisms  involved in functional MRI of the brain and to optimize imaging and data  analysis strategies for the detection of neuronal activity. Functional MRI  relies on the ability to detect the changes in NMR signal that are produced  in discrete regions of cortex in response to specific activating stimuli,  and are believed to reflect changes in local blood flow, volume and  oxygenation. Functional MRI promises to be a major addition to the methods  available for studying brain activation. Despite the widespread claims for  the power and successes of the method, there remain several unanswered  questions regarding its optimal mode of use, the tissue and technical  factors that are important in determining the signal changes detected, and  the significance and interpretation of these signal changes. The research  proposed would systematically address such issues. The underlying  mechanism may include both susceptibility contrast effects, based on the  BOLD effect, as well as wash-in effects, and these will be separately  quantified. The factors that affect each mechanism will be separately  identified and measured. For the BOLD effect, extensive computer modeling  and measurements in phantoms and animals brains will be used to establish  the relative sensitivity to vascular structures of different sizes,  spacings and orientations, as well as other tissue properties such as the  rate of water diffusion. The separate sensitivities to s-called static  field effects (T2\*), diffusive losses and other mechanisms will also be  established. The performance of different pulse sequences will be compared  to devise optimal methods of scanning and detection at 1.5T. Echo planar  imaging, conventional gradient echo and fast spin echo imaging as well as  more novel schemes will be compared in phantoms, animal brains and examples  of human activation. Human and animal activations will be produced in vivo  using visual and motor stimuli as well as by alteration of global blood  flow by acetazolamide and hypercarbia. A critical feature of current  paradigms for detecting activation is the method of data analysis, which is  interrelated with the nature of the task and imaging method used. We will  compare different methods of analyzing functional data sets, including  statistical parameter mapping, time-correlation analyses, and principal  component analysis. The sensitivity of each to motion and other artifacts  will be established by in in vivo comparisons and by computer simulations.  From these studies, we anticipate being able to improve strategies for the  use and interpretation of functional MRI in human studies of function and  cognition. | https://reporter.nih.gov/project-details/2714543 | R01 | 1998 | 1 |
| This proposal is directed toward improving tomographic imaging in  diagnostic radiology and nuclear medicine. It is predicated on the claim  that significant advances will be achieved in the fidelity of the images  that are reconstructed from the raw detector measurements of the  tomographic scanner by changing the basic elements (called "basis  functions") with which the image is built in the computer. The  conventional basic elements for computerized tomographic imaging are the  voxel basis functions, and the sinusoidal basis functions of Fourier  analysis. Two classes of promising new basis functions have been  developed: functions that are localized in space (as are the voxel basis  functions), and functions that are not localized (similar in many respects  to sinusoids). The new classes of basis functions are well-suited to  constructing faithful digital image representations of the biological  structures that have influenced the raw tomographic scanner data. The new  localized basis functions have a number of very desirable properties not  shared by voxels: they are rotationally symmetric, their Fourier  transforms are effectively localized, and they have continuous derivatives  of any desired order. The new non-localized basis functions are designed  to perform a spatially-variant filtering operation that is required by a  non-iterative method of 3D image reconstruction developed by the Principal  Investigator.    The specific aims are to develop mathematical theory, efficient computer  algorithms, application-specific implementations and evaluation criteria  for (1) methods of iterative reconstruction from projections, (2) methods  of estimating the fundamental limits on the performance of the  reconstruction process, and (3) methods of non-iterative 3D reconstruction  from projections. For specified imaging tasks, the level of statistical  significance will be found for rejection of the null hypothesis that two  methods perform a task equally well, in favor of the alternative  hypothesis that one method performs the task better.    The basis functions of the image representation are the essential core of  all methods for computerized image reconstruction, irrespective of the  medical imaging modality (e.g., CT, PET, SPECT, MRI). The development of  new computer algorithms and their associated image representations will  enable the full potential of scanners for functional imaging in emission  tomography (PET and SPECT) to be realized by extracting as much  information as possible from fully-3D low-statistics projection data. | https://reporter.nih.gov/project-details/2700454 | R01 | 1998 | 1 |
| DESCRIPTION: (Adapted from the Investigator's Abstract) The proposed research   has the goal of developing and validating interval psychometric scales of   visual function limitations and vision disabilities. These interval scales will   be developed using Rasch probablistic measurement models applied to ordinal   patient rating responses to individual questions. Once developed and validated,   these scales will be independent of the particular assessment used, as long as   the instrument is calibrated to the scale. The significance of the proposed   research is that it will provide a means of estimating measurements of latent   functional ability variables for individual patients with visual impairments.   In the future, these measurements can be used for parametric studies,   epidemiological studies, and clinical outcome studies. The proposed research   will identify the number and nature of functional ability scales. It will   determine the dependence of those scales on the diagnosis of visual system   disorder, the type of visual impairment, the existence of co-morbidities, and   the patient's history of rehabilitation. Existing visual function instruments   (NEI-VFQ, VF-14, ADVS, and VAQ) and two general function instrument's   individual items will be evaluated with respect to scales. To estimate the   scales, a large set of specific cognitive and motor activities (e.g., writing a   check) will be classified according to functional domain (reading, fine and   gross visual-motor, visual information processing [e.g., recognition,   localization, orientation], or mobility). In telephone interviews, low vision   patients will be asked to rate the difficulty of performing each activity.   Rasch analysis will be used to test the hypothesis that there is a global   functional ability scale and to test the validity of the a priori visual   function domains. Principal component analysis of response residuals will be   used to evaluate the dimensionality of visual function limitations. Patients   also will be asked to rate the difficulty of achieving specific activity goals   (e.g., cook a meal, manage personal finances) and Rasch analysis will be used   to estimate a vision disability scale. Item ordering and item intervals on the   scales and scale validity will be compared across diagnostic groups (AMD,   glaucoma, diabetic retinopathy, RP, CVA, and anterior segment disorders) and   for different types of visual impairments (e.g., acuity loss and contracted   visual fields). Person measures of functional ability will be evaluated as a   function of severity of visual impairments (visual acuity, contrast   sensitivity, visual fields, dark adaptation, color vision). Determining if the   NEI-VFQ, VF-14, ADVS, VAQ, SF-36, and SIP can be calibrated to common scales   will test the hypothesis that there is a common functional ability variable(s). | https://reporter.nih.gov/project-details/2911020 | R01 | 1999 | 1 |

**Cluster 22:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
|  | https://reporter.nih.gov/project-details/2230459 | R29 | 1995 | 0.81 |
| The Heart Failure Program project is a resource-related research program to  develop new reasoning methods for the application of Artifical Intelligence  techniques to medicine for the effort of the SUMEX-AIM community. The  context and driving force for this research is the management of heart  failure in the intensive care setting. We will: 1) Develop a  representational methodology capable of supporting the clinically relevant  distinctions of patient state. This representation will utilize the  clinically significant qualitative parameter values and will include causal  relationships, time dependencies and relations about change. 2) Build a  qualitative physiological model of the cardiovascular system using this  methodology to act as a store for evolving knowledge of patient state. 3)  Explore and develop strategies for determining the appropriate parameter  values in the model from input data, reasoning support methods and  heuristics for carrying out the diagnostic reasoning with the model, and  reasoning support methods and heuristics for carrying out the diagnostic  reasoning with the model, and reasoning support methods and heuristics for  finding possible therapies and determining their potential consequences.  4) Build around this core a program to assist the physician in exploring  his or her understanding of the implications in an individual case. The  physician and program will reason together about the case with the  physician providing ideas and the program assuring consistent consideration  of the implications. 5) Generalize the techniques for use in other medical  domains. | https://reporter.nih.gov/project-details/2217163 | R01 | 1996 | 0.81 |
| The elderly population is one of the fastest growing segments of the  American population. Nevertheless, there is a shortage of clinicians who  are specifically trained to identify and treat the problems associated with  advancing age. As a result, predictable deteriorations in renal function,  increasing drug-utilization over time and susceptibility to age-related  illnesses all serve to threaten the health of many elderly Americans. The  Geriatric Service of MediSource is a comprehensive computerized product  that addresses the informational needs of clinicians who care for elderly  patients. The Geriatric Service specifically offers age-specific  recommendations for preventative care, and tools to provide audits for  those medication-related hazards that mist commonly affect elderly  patients. The intent of the Geriatric Service is not to expand the content  of existing applications, but to create a fundamentally new approach to the  management of the elderly patient. We believe that integrated clinical  information services like those offered by Multum are rapidly becoming the  standard of care for pharmacologic and non-pharmacologic interventions and  hope to improve dramatically the care rendered to elderly patients with  this proposed Geriatric Service.    PROPOSED COMMERCIAL APPLICATION: The Geriatric Service of MediSource has  the potential to become the standard of care throughout the health care  industry for patient-specific drug information. The Geriatric Service can  be used by clinicians in ambulatory, inpatient, educational, and research  settings for current, concise, and patient-specific information related to  drug selection, dosing, and monitoring. | https://reporter.nih.gov/project-details/2002513 | R43 | 1997 | 0.81 |
| DESCRIPTION (Taken from application abstract): This proposed study will  replicate and extend methodology used in earlier studies and will use  extensive clinical data repositories, informatics tools, and expert  practitioners for perinatal medical knowledge building.    Clinical Data Repository: Duke University's Medical Center (DUMC) TMR (The  Medical Record) data repository will be used for this study, and contains  45,922 electronic medical records for both low and high-risk pregnant women  (and their infants) who have received prenatal care at DUMC, and its  affiliated regional clinics, between 1/1/86 and 12/3l/95. Each patient's  electronic data is used for clinical patient care and contains a potential  4000 variables per record. This volume of data requires new approaches for  data analysis and medical decision support, since human information  processing limitations become quickly overloaded by both an individual  patient s data and the aggregate information collected for the perinatal  patient population.    lnformatics Tools: Informatics techniques for knowledge acquisition and  data mining will use machine learning programs, statistical analysis, and  domain expert input to articulate relationships between the data and  perinatal patent outcomes. The goal is to provide decision support for  perinatal care providers to accurately identify patients at risk and assist  them with modifiable preterm birth ask factors. An expert system will use  data-generated and verified knowledge bases to test its predictive validity  when new patient cases are induced to the expert system. Earlier studies  found 53-90% predictive accuracies for an expert system prototype, as  compared to 17-38% accuracies, reported in the literature, using current  manual techniques. Mapping the expert system's knowledge base terms to  medical library resources will be explored for additional decision support.    Expert Practitioner: The perinatal expert panel will consist of the  Principal Investigator, a Board Certified OB-Gyn Physician, and a certified  Perinatal RN. Each of the panel members has more than 20 years of perinatal  experience. Participating informatics experts are known, both nationally  and internationally for their expertise in the field of Medical Informatics. | https://reporter.nih.gov/project-details/2714219 | R01 | 1998 | 0.71 |
| Our long-range objective is to understand the functional organization  and dynamical activity of the cortex. The discovery of the columnar  organization of the cortex has led to the notion that the columns are  fundamental building blocks, from which larger functional units are  constructed. The cortex is thus viewed as a crystal (a more or less  regular array of repeating, similar modules. Our proposal will test and  refine this modular hypothesis.    We shall use optical imaging of the primary visual cortex of monkeys and  cats, and simultaneously record electrical responses from small neuronal  clusters and local field potentials. We shall thus obtain a spatio-  temporal picture of the activity in the neural ensembles which encode  various stimulus parameters. The data will be analyzed with extensions  of Principal Component Analysis that we have developed.    We address three major aims: 1) To test the modularity hypothesis we  shall measure, in a large piece of cortical tissue, the full range of  functional maps ( for orientation, color, spatial frequency etc.)  together with the retinotopic map. We shall measure the periodicity of,  and correlations among, the functional maps, to determine if they are  commensurate. This will lead to a refined framework that could include  possibly incommensurate cortical scales and interactions among cortical  elements. 2) We shall investigate how the Principal Components  (eigenfunctions) obtained from the optical images depend on the extent  of the visual stimulus, to determine how the dynamical dimension of the  primary visual cortex (viewed as a dynamical system) scales with size.  3) We shall study the concerted electrical responses of neuronal  clusters, to clarify the link between optical signals and neuronal  activity, and to deepen our understanding of the neuronal dynamics.    Our study is aimed at an intermediate architectural level, and deals  with the way in which the fundamental modalities of the visual world  (orientation, size, color and so on) are analyzed in the primary visual  cortex. Such knowledge is crucial for the construction of cortical  models, which are essential for any quantitative understanding of  critical function and dysfunction. | https://reporter.nih.gov/project-details/2890529 | R01 | 1999 | 0.71 |

**Cluster 23:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| New prosthetic methods are giving people with motor impairments  alternative communication and control channels. A logical culmination of  these developments is a system that allows the brain to bypass completely  its normal output pathways. Recent studies from this laboratory have  shown that humans, including those with motor disabilities, can learn to  change rapidly and accurately the amplitude of the 8-12 Hz mu rhythm in  the electroencephalogram (EEG) recorded over sensorimotor cortex.  Furthermore, they can use this control to move a cursor on a computer  screen. Good single-channel control has been obtained, and initial data  indicate that multichannel control is also possible. Thus, the mu  rhythm, which recent work shows is detectable in nearly all adults, may  support a multichannel brain-to-computer interface, and thereby provide a  powerful new communication and control option for severely disabled  individuals.    This project's goal is a reliable multichannel brain-computer interface.  The proposed approach is based on three well-supported hypotheses: that  the scalp-recorded 8-12 Hz mu rhythm comprises a number of relatively  independent components, that topographic analysis and frequency analysis  techniques can distinguish between these components, and that humans can  learn to control specific components and use them to operate a  multichannel brain-computer interface.    The first objective is to define by topography and frequency the separate  8-12 Hz components that are present when individuals are using the  current interface. The second objective is to determine which components  individuals are best able to control. The third objective is to  incorporate these trainable components into a multichannel brain-computer  interface that is rapid and reliable. These objectives will be achieved  by combining online studies in which subjects learn to use the interface  while extensive data are stored, and offline data analyses in which  methods for improving the interface are defined.    This project should produce an EEG-based brain-computer interface of  significant value to individuals with disabilities. It should also lead  to further work exploring the practical capabilities and theoretical  implications of this new form of communication. | https://reporter.nih.gov/project-details/2202489 | R01 | 1995 | 1 |
|  | https://reporter.nih.gov/project-details/2237592 | F37 | 1996 | 1 |
| Efforts to apply computer methods to assess and improve the quality of  care in the hospital have been stymied by limited access to clinical  data. Free-text data have detailed clinical descriptions of patients  that would be useful in computer altering systems and computer reminder  systems. However, free-text data cannot be interpreted by most clinical  computer systems. In this proposal, we describe research specifically  aimed at making free-text data accessible to computer-based applications  for assessing and improving the quality of care. In particular the  research plan focuses on the development of technologies that would allow  free-text data to be used in clinical alert systems for critical test  results; in reminder systems to encourage adherence to practice  guidelines; and in data collection systems for severity of illness models  applied in the assessment of risk adjusted outcomes. The approach  described in the research plan emphasizes the development of statistical  and probabilistic methods for interpretation of data derived from medical  language processing systems. We will test the methods developed for  language processing and interpretation developed under this proposal in  3 area: 1) the identification of concepts related to severity from the  MedisGroups and the Computerized Severity Index models of patient  severity of illness; 2) the identification of chest x ray reports and  mammography reports with potentially malignant findings that require  radiological follow-up; 3) and the automatic assessment of  appropriateness of coronary artery bypass grafting (CABG) surgery from  free-text descriptions of patients based on the application of a clinical  practice guideline for CABG surgery. | https://reporter.nih.gov/project-details/2430870 | R29 | 1997 | 1 |
| The PI is a Neurophysiologist and Pediatric Neurosurgeon who will be the  Associate Director of Center V (Behavioral and Neurobiology Research), with  special responsibility for Neurobiology programs, at the Children"s  Research Institute. This RCA will permit the PI to devote nearly full time  to research.    Short term goals are to test the hypothesis that neuronal ensembles have  nonlinear deterministic properties. If so, they will 1) have activity that  can be characterized and controlled through unstable periodic orbits, 2)  hen noise driven will exhibit stochastic resonance, and 3) because of  coupling will exhibit generalized (nonlinear) synchrony and emergence.  Long-term goals are to achieve a better understanding of neuronal network  and brain behavior, and to develop novel methods of treating dynamical  diseases.    The research project will involve theoretical work on the detection of  unstable orbits in In Vitro brain slices and human epileptic foci. Such  orbit information forms a novel method of characterizing the deterministic  properties of complex systems despite nonstationarity, and can be used to  control those systems. Nonlinear systems also optimize their response to  weak signals in the presence of noise - stochastic resonance. We will  define the statistical mechanics of stochastic resonance through  simultaneous measurements of single neuron and neuronal ensemble activity.  Since neuronal ensembles may demonstrate nonlinear generalized synchrony,  we will quantify spatio-temporal generalized synchrony through dual  simultaneous single cell recordings as a function of separation in a  neuronal network. Both unstable orbit detection and generalized synchrony  will be used to define the emergence of nonlinear behaviors in neuronal  ensembles.    The results of this research will fundamentally alter the way that neuronal  dynamics can be characterized and controlled, will provide a means to deal  with neuronal nonstationarity, will further explain the role of noise in  the nervous system, and may provide a novel approach for the control of  pathological neuronal ensembles in dynamical diseases such as epilepsy,  spasticity, and tremor. | https://reporter.nih.gov/project-details/2674504 | K02 | 1998 | 1 |
|  |  |  |  |  |

**Cluster 24:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| DESCRIPTION: (Adapted from the applicant's abstract): Medical and  biological data often come in the form of digitized signals and images, for  example magnetic resonance images (MRI), ion channel electrical series, and  human gait paths. As data acquisition becomes easier, sequences of such  images or signals are collected, often along with other covariate  measurements, resulting in data sets where the basic unit of measurement or  response is a high dimensional object. This project proposes a battery of  statistical techniques for modeling and understanding such data, that  explicitly takes into account and indeed exploits the inherent, spatial, or  temporal correlation, and when appropriate, relates it to covariate  information. By imposing spatial smoothness in the image or signal domain,  pixel-wise regression, and canonical correlation models can borrow strength  from neighboring pixels. This not only improves the overall efficiency of  these techniques, but also allows identification of important regions rather  than individual pixels. The project develops appropriate versions of  nonparametric regressions for such series of images, as well as data  descriptions such as clustering, principal component, and singular value  decomposition models. In many cases, wavelets will be used to achieve  spatial smoothness. In the case of ion channel data, the models are used to  isolate particular weak high frequency components from correlated noise.  Much of this work will be carried out in collaboration with radiologists,  physiologists, and other biomedical researchers working on cancer, heart  disease and stroke, brain mapping, and gait analysis. | https://reporter.nih.gov/project-details/2517753 | R01 | 1997 | 0.82 |
|  | https://reporter.nih.gov/project-details/2638556 | R29 | 1997 | 0.82 |
| The goal of this research is to develop improved assessment protocols  that afford a quantitative and analytic evaluation of speech impairment  in children and adults with neurological disorders. Speech impairments  (dysarthrias) will be studied in persons with amyotrophic lateral  sclerosis, Parkinson's disease, stroke, cerebellar degeneration, cerebral  palsy, and developmental speech disorders. Improved evaluation of speech  intelligibility is a particular focus of this work, but issues of speech  and voice quality also are addressed. The methods to be used are a  combination of standard clinical assessments (such as rating scales),  intelligibility evaluations, and computer-based acoustic analyses.  Specifically, the methods include: perceptual ratings of speech by  experienced clinicians, quantitative assessment of intelligibility, a  multiple-parameter acoustic analysis, computer correction of speech  abnormalities through LPC resynthesis of the acoustic signal, and  derivation of vocal tract shape from acoustic parameters. Work in all  these areas will be based on recordings of speech samples from a large  number of individuals with dysarthria. One product of the research will  be a library of clinical profiles including intelligibility scores,  phonetic feature analyses, ratings of speech/voice quality, acoustic  measures, and neurological diagnosis. Particular attention will be given  to the influences of subject age and sex on the characteristics of  dysarthria for a given neurological diagnosis. The assessment protocols  will be implemented on microcomputers and designed to be incorporated in  clinical practice. The research also will contribute to the development  of expert systems for the rating and classification of dysarthria. | https://reporter.nih.gov/project-details/2733645 | R01 | 1998 | 0.82 |
| The proposed mentored research award aims to develop the applicant's  basic research skills and to prepare her for independent  investigations directed at exploring and dissecting the neural bases of  motor and related dysfunctions in neurological disorders and normal  aging using imaging techniques, psychophysical studies and brain  network modeling. The primary mentor for this plan is Dr. David  Eidelberg, a neurologist with vast experience in the field of PET  imaging and movement disorders. Dr. Claude Ghez, a neuroscientist  who has worked in the field of motor control for over 25 years will  serve as co-mentor. Dr. Glyn Johnson, experienced physicist in the  field of magnetic resonance, will serve as consultant to instruct the  applicant in functional MRI. Drs. Jim Moeller and Eva Petkova will  further the applicant's education in advanced statistics. Dr. Vijay  Dhawan will be consultant for PET biophysics and quantification. Dr.  Ken Perrine, neuropsychologist, will also be available as consultant.    Research plan: The goal of this study is to characterize changes in  neural processing underlying implicit and explicit forms of motor  learning in patients with Parkinson's disease and in a control normal  aging population. A new family of motor tasks developed in Dr. Ghez's  lab will be used during 150-H2O positron emission tomography. A new  analytical methodology based on principal component analysis will be  applied on measures of regional blood flow to determine how inter-  subject differences in motor learning are reflected in modulations of  brain network expression. These same methods will be used to  determine the neural bases of two therapeutic inventions in  Parkinson's disease, ventral pallidotomy and pallidal stimulation, and  to assess their effectiveness in reversing alterations in motor  learning.    Educational plan: The educational plan will focus on the following  objectives: 1) Completion of the research plan. 2) Learning imaging  techniques, including image acquisition and analysis using PET and  fMRI. 3) Development of testing techniques to be used in fMRI. 4)  Acquisition of skills in brain network analysis (including SSM) to be  applied to both PET and NMR imaging. 5) Furthering the applicants  statistical and computational skills through courses taken at Columbia  University School of Public Health. 6) Developing the applicant's  laboratory management, supervision, research communication and  mentoring skills. 7) Receiving instruction in the responsible conduct  of research. | https://reporter.nih.gov/project-details/2891436 | K08 | 1999 | 0.82 |
| The genes of the human leukocyte antigen (HLA) region control a variety  Of functions involved in the immune response, and influence  susceptibility to over 40 diseases. Our understanding of the structure  and function of the HLA genes, their disease associations, and the  evolutionary features of this multigene family has benefitted from recent  advances in molecular biology, immunology, disease modelling and  population genetics. Theoretical studies in the development of models to  determine the modes of inheritance of the HLA associated diseases have  led to a better understanding of the inheritance patterns in insulin  dependent diabetes mellitus, rheumatoid arthritis, multiple sclerosis,  ankylosing spondylitis, hemochromatosis, celiac disease, and others. It  is now clear that many of the HLA associated diseases involve  heterogeneity in their HLA components, as well as non-HLA genetic  components.    The specific aims of our research are to study the genetic components in  the etiology of the HLA associated diseases, and population genetic  features of the HLA system. A variety of methods to test modes of  inheritance of diseases using marker allele information, will be  developed. Methods appropriate for the analysis of marker systems which  are not highly polymorphic, to both detect linkage and determine modes of  inheritance, will be investigated. The information content of particular  pedigree types for LOD score analysis will be investigated. Two methods  using patterns of linkage disequilibrium will be investigated to  determine their usefulness in mapping disease predisposing genes. A  number of large collaborative data sets of HLA associated diseases will  be analyzed. A framework for genetic counselling of HLA associated, and  other complex diseases, will be developed. The results of our studies  are generally applicable to the mapping and characterization of complex  human genetic traits. | https://reporter.nih.gov/project-details/2196932 | R01 | 1995 | 0.36 |

**Cluster 25:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| Therapeutic management based on protocols, and clinical algorithms is  increasing in significance as medicine tackles more serious illnesses and  must do so in cost-efficient manner. Clinical algorithms are appealing as  they provide a good overview of the clinical state of a patient, identify  key decisions and the outcomes associated with those decisions. They  provide guidance and improve the quality of care without taking the  control away from the physicians. Despite these advantages clinical  algorithm design, specification, and delivery remain problematic. This  research proposal seeks to overcome these difficulties, while retaining  the advantages of clinical algorithms.    We propose to develop a general framework for the representation for  clinical algorithms and supporting medical knowledge, to communicate these  algorithms effectively by extracting and presenting protocols specific  individual patient's condition, and to provide decision aids for protocol-  based patient care. This framework will use knowledge about therapy in the  context of specific conditions to create patient-specific algorithms  refined from the general-purpose algorithms. Aspects not pertinent to the  case will be removed. The derivation of the specialized algorithms from  the general algorithm will be recorded and used to explain the logic of  the specialized algorithms. Because the algorithms will be tailored to the  patient context, they will be detailed, yet manageable. This will avoid  the over-simplifications of current "one-size-fits-all" algorithm. Mixed  graphic and textual presentation techniques will be used to develop a  user-friendly interface with explanation facilities integrated with  patient information systems. We will test the framework by developing  operational systems for hemodynamic resuscitation and management of  circulatory shock in critically ill postoperative patients.    We have chosen the domain of the management of circulatory failure in  critically ill postoperative patients because: (a) circulatory failure is  the significant medical problem affecting many postoperative patients  where proper management can mean the difference between life and death,  (b) we have a long track record of on-going activities in development of  objective treatment methodologies in this area, (c) we have developed and  tested a general clinical algorithm for management of high risk  postoperative surgical patients, (d) our recent attempts to extend and  refine the initial algorithm using traditional approaches have been  frustrated by the complexity of the task. The proposed framework will  support the creation of detailed therapeutic management algorithms and  decision support systems to implement these algorithms in daily practice. | https://reporter.nih.gov/project-details/2237763 | R01 | 1995 | 0.83 |
| Neural network optimization algorithms greatly enhance our ability to  construct large-scale, dynamical models of highly interconnected networks.  Until now, optimization has only been applied to networks of simplistic  processing units, ignoring the integrative and temporal response  properties of single neurons, thus limiting the predictive power of the  models. The long-term goal of this project is to develop a hybrid  modeling strategy in which optimization methods are applied to networks of  realistic,multicompartmental model neurons. To accomplish this goal, we  will construct a hybrid model of an actual distributed processing network  composed of repeatably identifiable sensory, motor, and interneurons that  computes a well-defined behavioral input-output function. Optimization  will be used to predict the connectivity of as-yet-unidentified  interneurons in the actual network and the predictions will be tested by  identifying the interneurons by physiological and morphological means.  Performance of the hybrid model will be assessed by comparing it to the  performance of an a priori model in which all connection strengths are  determined physiologically. The final model will be used to predict the  loci of synaptic plasticity underlying nonassociative conditioning of the  reflex by incorporating local learning rules and by optimization methods.  The predictions will be tested by determining the actual plastic sites  physiologically. This project will have the combined effect of enhancing  the predictive power of optimized network models and illuminating the  relation between computations at the single-neuron and network levels. | https://reporter.nih.gov/project-details/2416012 | R29 | 1997 | 0.83 |
| DESCRIPTION: Limitations of many disabled individuals' production efficiency, communication rate and vocabulary organization impact negatively on their ability to communicate competently using present-day Augmentative and Alternative Communication (AAC) devices. This project proposes a re-conceptualization in the way language units are organized and retrieved in present-day AAC technologies, by operationalizing recent linguistic research in frame theory. A frame-based organization for utterances is proposed by exploiting natural category systems found in the context, culture and conceptual structures of the individual. A method and system for implementing these knowledge structures on a computer system will be developed. Specifically, the purpose of Phase l of this project is to provide a proof of concept for the Frametalker technology design. During the course of the year we will a) develop a software prototype of the Frametalker, b) integrate it with a database of communication frame-structured utterances and c) submit Frametalker to laboratory and field-testing. | https://reporter.nih.gov/project-details/2850474 | R43 | 1998 | 0.83 |
| The brain is at risk of serious cerebrovascular insult during the  400,000 cardiopulmonary bypass (CPB) surgeries performed annually in the  United States. There is strong clinical evidence that over two-thirds  of patients exhibit neurologic or neuropsychologic (NP) Postoperative  deficits caused by emboli passing to the brain during surgery. There  is a critical need for a device that can not only detect emboli, but  also classify them as to type. Classification is critical for  correlating neurological deficits with emboli type, determining the  source of the emboli, and subsequently changing surgical procedures  and/or administering neuroprotective agents to minimize brain injury.  Continuous-wave Doppler ultrasound equipment can detect emboli, but  cannot provide the information needed to classify emboli composition.  In Phase I, ORINCON and the Bowman Gray School of Medicine utilized  broadband pulse-echo ultrasound and an artificial neural network to  demonstrate significant classification capability on in-vitro data. In  Phase II, we will refine and integrate components from Phase I into a  PC-based system capable of accurate, real-time emboli detection and  classification in both extracorporeal pump circuits and the carotid  artery. Extensive in-vitro and in-vivo data will be collected to permit  refinement and thorough evaluation of classification capabilities.  Clinical studies will be performed to correlate neuropsychologic  deficits with embolus composition.    PROPOSED COMMERCIAL APPLICATION:  There is immediate commercial potential for a low-cost (less than 25K),  real-time automated emboli detection and classification system. The  market includes manufacturers of cardiopulmonary pump circuit devices  (Medtronics, Pall, Cobe, Sarns), hundreds of medical centers performing  cardiopulmonary bypass, and producers of neuroprotective drugs (Astra,  Bayer, Sterling). A letter from Medtronics expressing this commercial  potential is enclosed. | https://reporter.nih.gov/project-details/2891987 | R44 | 1999 | 0.83 |
| The first project objective is directed to developing a comprehensive  computer interactive neuropsychologic test battery that is sensitive for  detecting subtle disturbances of cerebral integrity. This task will be  accomplished accessing a panel of neuropsychologists, an expert systems  specialist and a software architect. Next, standardization will be  conducted on 500 normal adolescents and adults. This sample will be  stratified by age and gender and demographically representative of the  population. Concurrent to these activities, an expert system will be  created. This latter task will enable dissemination of the final  protocol to the widest possible user audience. Finally, the interactive  test battery will be used to ascertain the prevalence, types and severity  of neuropsychologic deficit in samples of adolescent and adult drug  abusers.    The ultimate goal of this research program is to devise a standardized,  quantitative and comprehensive method for characterizing the cognitive  and psychomotor capacities of drug abusers. Computer interactive  testing, because of its precision in measurement, will enable the  detection of subtle or occult impairment residual to chronic drug abuse.  For this reason, the battery is to be created specifically for this  population. The information obtained from such an assessment has  potentially important ramifications for determining the timing of  treatment, type of treatment and post-treatment vocational rehabilitation  of drug abusers. | https://reporter.nih.gov/project-details/2119052 | R01 | 1994 | 0.34 |

**Cluster 26:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The goal of the proposed research is the analysis of biological sequence  data to address the molecular mechanisms of evolution and the origin(s)  of all viruses and related genetic elements. Phylogenetic trees will  provide a framework for the mapping of cell and tissue tropism,  pathogenicity and virulence, modes of transmission and geographical  distributions, and many other higher order characteristics of viruses.  The specific aims of proposed analytical studies are: i) determining  functionally equivalent networks and frequency of exchange among and  between retroid elements, and their potential cellular homologues,  including new studies on 300 retroviral env proteins; 2) inferring  functionally important regions of all proteins of paramyxo-, rhabdo- and  filoviruses, (with privileged access to new Ebola sequences), and Borna  Disease virus, (including potential BDV sequences from schizophrenic  patients); and 3) the analysis of the dUTPase gene, as a model system,  to address issues relevant to the structure, function and evolution of  duplicated sequences, and potential horizontal transfer among and between  host and viral genomes. The specific aims of the technical studies are:  i) evaluation of stochastic production model approaches for generation  of multiple alignments, detection of recombination, and calculation of  evolutionary distances; and 2) development and testing of new and  existing methods for historical reconstruction of functionally equivalent  networks.    RNA viruses (e.g. HIV, or Ebola) are the major causative agents of human,  animal and plant viral diseases world wide. The heterogeneous nature of  RNA populations makes it difficult to develop effective, anti-viral  agents. The sequence database is now large enough to conduct comparative  studies on natural variants versus chemotherapeutically induced mutants  for several retroviral proteins. This model study will provide new  information on the nature of selected mutations which will be useful in  future anti-viral drug development.    Computational analysis of primary sequence data is an area of intense  interest in biology, mathematics, statistics and systems science. In the  last few years new approaches to problem solving and classification, such  as machine learning, neural networks, genetic algorithms, and stochastic  production models or, "intelligent systems" as they are referred to  collectively, have become available. Unfortunately most biologists are  unaware of these developments. Application of these methods to real data  remains unexplored. The proposed studies will go a long way in rectifying  this gap in technological utilization. These studies will continue to  define important evolutionary' relationships and events, provide  biologically informative sequence relationships for bench-marking new  software, and contribute new information relevant to the structure and  function of viral proteins suggesting new directions in laboratory  experimentation. Strategies and techniques developed for the analysis of  highly divergent genomes can also be applied to the study of the wealth  of sequence information generated under the auspices of the Human Genome  Project. | https://reporter.nih.gov/project-details/2057512 | K04 | 1995 | 0.73 |
| DESCRIPTION: Over the past decade a variety of alternative computer  based modeling techniques have been introduced which show  promise for the construction of clinical decision aids. These  techniques include statistical regression approaches such as  generalized additive modeling, classification tree induction such as  ID3 or CART, and multi-layer neural networks. Logistic regression  models (LR) are currently central to most probabilistic predictive  clinical decision aids and are fundamental to comparative analyses of  medical care based risk adjusted events. These newer techniques  have been applied on a larger scale in the last few years. They  appear to have unique advantages in selected circumstances. The  successful use of these methods, however, depends on understanding  their accuracy, performance, and model transportability.    A formal assessment of these new techniques with four specific aims  is proposed: (1) to assess and compare the performance of  different models to determine the factors which affect  performance; (2) to develop automated computer based procedures  for exploratory model development for each method; (3) to develop  hybrid models incorporating the strengths of each of the existing  techniques, and (4) to determine the situations that restrict the  transportability of these models.    These specific aims will be achieved in a three stage project. In  the first stage four approaches will be pursued: (1) the  mathematical properties of the different computational algorithms  for the modeling techniques will be studied; (2) automated  modeling procedures will be developed and utilized; (3) the factors  that affect performance for each modeling technique will be  explored and(4) new hybrid techniques will be developed and assessed.  In the second stage the methods developed in the first stage will  be used to create and test models that predict cardiovascular  events on data from 15,000 patients in a prospective clinical  trial. In the third stage the factors that affect the  generalizability and transportability of models to new datasets will  be explored by repeated sampling and model construction on  different subsets of the cardiovascular database including  separating the database into subsets from each of ten different  hospitals.    This work will broaden the understanding of these important  modeling techniques and their potential contributions for  clinical decision making, health policy research, and medical  informatics. New modeling techniques might be developed which  incorporate elements from different techniques. | https://reporter.nih.gov/project-details/2032352 | R01 | 1997 | 0.73 |
| The overall goal of this project is to develop a comprehensive computer  model of neural coding of auditory space in the auditory thalamocortical  system. A self-organizing neural network model of auditory cortical maps  will be studied. The output of the network will be that of a simulated  planar cortex that will afford direct comparisons of simulated virtual  space receptive fields (VSRF) with those of actual VSRFs of direction-  sensitive neurons in primary auditory cortex (AI). These computational  simulations will greatly benefit from the ability to synthesize virtual  auditory space from measured Head-Related Transfer Functions (HRTFs). The  same stimuli used for microelectrode recordings in auditory cortex of cat  will be used in the development and stimulation of the neural-network  models. These network models assume that primary auditory cortex is  subject to experience-dependent changes. Currently available  computational models of neural signal processing in the auditory periphery  and brain stem will be used to provide a neural representation of binaural  stimuli to the self-organizing thalamocortical model. The simulation of  self-organizing processes operating on input data with intrinsic structure  leads to the emergence of topographical maps. These maps afford the  opportunity to examine overlays of functional organization. Currently,  the ability to corroborate the emergence of a spatial auditory map with a  detailed map of an auditory cortical field is quite limited, but the  simulated development of maps will demonstrate how global topographic  order can emerge, in principle, from local cooperative and competitive  interactions within the cortical field. It is anticipated that these  simulation studies will help guide neurophysiology research with regard to  deciphering the neural code of auditory space. Specifically, the  simulations may suggest where to probe the cortex with microelectrodes and  with what types of stimulation. Interactions between the tonotopic  frequency organization and orthogonal iso-frequency organization will be  investigated computationally. This computational modeling work may  provide a better understanding of the representation of complex sounds in  general at higher levels in the auditory system. Given the nature of the  model to re-organize, effects of cochlear lesions can be studied and thus  aid in the study of sensorineural hearing impairment. | https://reporter.nih.gov/project-details/2414667 | R03 | 1997 | 0.73 |
| Two-dimensional polyacrylamide gel electrophoresis (2DGE) can detect  thousands of polypeptides, separating them by apparent molecular weight  and isoelectric point. It provides a more realistic and global review  of cellular genetic expression than any other technique. Computer 2DGE  analysis software for a variety of platforms has been developed to deal  with what can easily be an unwieldy data management problem. However,  these programs require significant user input for gel matching and do  not adequately correct for routine gel-to-gel distortions. The best  approach for gel matching and spot detection is pixel matching from one  gel to another. This procedure is computationally demanding and is  currently limited to high-end workstations costing 80,000 dollars and  above. This Phase I effort proposes to utilize a pixel matching  algorithm and integrate an enhanced image processor that will fully  automate spot detection for a PC-based system. This system will perform  at the level of the expensive workstations but cost 40-50 percent less.  Phase I goals include hardware integration of a new image processor PC  card into a Pentium class computer, software development for pixel  matching of reference to study gel and preliminary testing of the  software to meet expected benchmark processing speeds for 2D plasma  protein maps. Full automation and increased performance in a PC-based  product will facilitate the use of 2DGE in the clinical setting and in  proteome research where the focus is on understanding the order,  regulation, and coordination of the human genome.    PROPOSED COMMERCIAL APPLICATION  The proposed project will result in a PC-based system capable of fully  automated spot detection for 2D gel analysis. The algorithm used will  correct for gel-to-gel variations without significant user input. This  affordable system will speed up 2D gel analysis and will allow for wider  clinical use of this valuable analytical technique. | https://reporter.nih.gov/project-details/2536786 | R43 | 1998 | 0.73 |
| DESCRIPTION (Taken from application abstract): This proposed study will  replicate and extend methodology used in earlier studies and will use  extensive clinical data repositories, informatics tools, and expert  practitioners for perinatal medical knowledge building.    Clinical Data Repository: Duke University's Medical Center (DUMC) TMR (The  Medical Record) data repository will be used for this study, and contains  45,922 electronic medical records for both low and high-risk pregnant women  (and their infants) who have received prenatal care at DUMC, and its  affiliated regional clinics, between 1/1/86 and 12/3l/95. Each patient's  electronic data is used for clinical patient care and contains a potential  4000 variables per record. This volume of data requires new approaches for  data analysis and medical decision support, since human information  processing limitations become quickly overloaded by both an individual  patient s data and the aggregate information collected for the perinatal  patient population.    lnformatics Tools: Informatics techniques for knowledge acquisition and  data mining will use machine learning programs, statistical analysis, and  domain expert input to articulate relationships between the data and  perinatal patent outcomes. The goal is to provide decision support for  perinatal care providers to accurately identify patients at risk and assist  them with modifiable preterm birth ask factors. An expert system will use  data-generated and verified knowledge bases to test its predictive validity  when new patient cases are induced to the expert system. Earlier studies  found 53-90% predictive accuracies for an expert system prototype, as  compared to 17-38% accuracies, reported in the literature, using current  manual techniques. Mapping the expert system's knowledge base terms to  medical library resources will be explored for additional decision support.    Expert Practitioner: The perinatal expert panel will consist of the  Principal Investigator, a Board Certified OB-Gyn Physician, and a certified  Perinatal RN. Each of the panel members has more than 20 years of perinatal  experience. Participating informatics experts are known, both nationally  and internationally for their expertise in the field of Medical Informatics. | https://reporter.nih.gov/project-details/2032587 | R01 | 1997 | 0.44 |

**Cluster 27:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The end goal of research in natural language processing is to provide a  psychologically plausible computational account of human language  processing. To this end, cross-linguistic investigation of sentence  processing can provide critically significant results which elucidate the  interaction between the grammars of particular languages and the  mechanisms of human language processing system. In the proposed research,  one critical aspect of sentence processing, namely, how speakers of  Japanese and English deal with temporal ambiguities that are encountered  during on-line processing of a sentence.    In particular, we ask how speakers of Japanese deal with a high degree of  indeterminacy early in the sentence. In order to investigate whether  Japanese speakers are engaged in extensive reanalyses en route to a  correct interpretation, five types of sentences are proposed to be  examined: sentences with relative clauses; sentences with stative verbs;  passive and causative sentences; sentences with verbs of giving and  receiving; and sentences with lexical homonymy. These sentences all  contain structures that might trigger reanalysis. Syntactic reanalysis of  grammatical relations as well as reanalysis of thematic role assignments  will be examined. The sentences will be tested by three psycholinguistic  experiments: eye-movement monitoring; self-paced reading methods; and on-  line grammaticality judgement task. Two questionnaires concerning peoples'  subjective judgments of conscious reanalysis and difficulty of the  sentence will also be conducted to supplement the experimental data.    The results of the Japanese studies will then be contrasted to English  data (available in the literature) to ask whether the ways in which  Japanese speakers process temporal ambiguity are similar to that of  English speakers. Using these cross-linguistic results, we will address  the question of whether a single mechanism can account for the way  speakers of Japanese and English deal with temporal ambiguity, or  alternatively, whether their differences are better accounted for by  invoking two different parsing mechanisms. | https://reporter.nih.gov/project-details/2445534 | R29 | 1997 | 0.79 |
| DESCRIPTION (adapted from applicant's abstract): A central issue in  behavioral neuroscience is how alterations in neural pathways mediate the  durable behavior changes involved in learning. Taste aversion conditioning  is an excellent model for studying the neural changes involved in learning  because this conditioning can occur in a single trial, despite lengthy  delays between conditioned and unconditioned stimuli. The proposed studies  are based on the identification of a cellular correlate of the behavioral  expression of a conditioned taste aversion, namely, cells in the nucleus of  the solitary tract. C-fos induction occurs in cells in the nucleus of the  solitary tract in response to a taste made aversive by conditioning, but not  in response to the same taste prior to conditioning or to a taste (quinine)  which is innately aversive. Proposed studies will combine this cellular  measure with behavioral assessment to further assess the cellular c-fos  response and its reliability as a marker of learning. Studies will also  examine the functional importance of cells in the nucleus of the solitary  tract which display c-fos induction during expression of this learning,  using asymmetrical lesion techniques. Additionally, studies will continue  to define the forebrain pathways critical to this learning, with a focus on  ipsilateral connections between the amygdala and the nucleus of the solitary  tract which appear necessary for cellular, as well as behavioral,  manifestations of this learning. The role of insular cortex will also be  examined. Finally, studies will identify the targets of activated modified  behavioral response to a taste after conditioning. By defining the neural  pathways and cell types involved in CTA learning, this project provide the  groundwork for eventually characterizing the plastic changes within and  between cells which underlie this learning. | https://reporter.nih.gov/project-details/2450926 | R01 | 1998 | 0.79 |
| LONG-TERM OBJECTIVES 1. Develop a computerized system, based on  hierarchical neural network pattern recognition technology, for reliable  identification of plants. 2. Identify poisonous plants. 3. Expedite  discovery of new medicinal plants. 4. Create an image database directly  from plant material and link with existing medicinal plant databases.  5. Develop commercial product for pharmaceutical companies, agriculture  and others. SPECIFIC AIMS 1. Design hierarchical system of neural  networks to follow natural plant taxonomy groupings and extend our  identification technology to a large number of plant species. 2.  Improve accuracy of identification. 3. Design a prototype workstation  for botanical and agricultural field stations and laboratories.    RESEARCH DESIGN AND METHODS FOR ACHIEVING GOALS. 1. Digitize large  number of plant species from special collections. 2. Measure  automatically venation patterns and shape. 3. Design hierarchical  neural networks to divide plants into natural groupings. 4. Accumulate  virtual herbarium database as leaves are digitized (scanned or  photographed).    POTENTIAL FOR TECHNOLOGICAL INNOVATION This system is unique in  capturing botanical recognition knowledge in a hierarchy of neural  networks and is the first fully-computerized system for plant  identification utilizing information digitized directly from plants.    PROPOSED COMMERCIAL APPLICATION  1. Expedite discovery of new medicinal plants for pharmaceutical  industry. 2. Create valuable database directly from plants. 3.  Identification of poisonous plants. 4. Valuable for rapid  identification of invasive weeds. | https://reporter.nih.gov/project-details/2895286 | R44 | 1999 | 0.77 |
| This application is a request for a NIH Research Scientist Development  Award, Level II RSDA-II) to extend work supported under the RSDA-I  previously awarded to the applicant (DA00139). (The RSDA-I granted in 1989  was titled "Innovative Statistical Approaches to Drug Abuse Data"; this  application for the RSDA-II is titled "Drug Abuse: Epidemiology, Treatment  Processes, and Outcomes.") The RSDA-Il will continue to ensure financial  stability and release time from the pursuit of funding for actual research  work. The major focus for the applicant during this five-year award is  continuation of her professional work examining drug use epidemiology and  treatment interventions for problematic drug abuse. Examining the  implications of research findings for treatment strategies and developing  the necessary social policy changes to support the implementation of  improved treatment strategies is of further interest. The applicant will  continue her professional work applying innovative statistical  methodologies to drug abuse data. To this end, three convergent lines of  current research will be continued. The first examines drug use and  treatment utilization among subjects recruited through hospital emergency  rooms sexually transmitted disease clinics, and jails. The second project  is to improve the efficacy and efficiency of matching drug users' treatment  needs to services. The third involves examining and evaluating the process  of treatment service delivery with a special focus on the roles and  functions of drug treatment counselors.    The applicant's supporting institution is a research unit, the  Neuropsychiatric Institute (NPI), organized within the Department of  Psychiatry, School of Medicine, UCLA. Affiliated with the NPI is the UCLA  Drug Abuse Research Center, which has been conducting research in drug  abuse epidemiology natural history of narcotics addiction, treatment  evaluation, and social policy over the past 20 years. In this setting, the  applicant will conduct the proposed research and will receive additional  training in psychiatric aspects of drug abuse treatment and in the  implementation of treatment services. Furthermore. the applicant's  considerable psychosocial research knowledge and skills in drug abuse  issues will contribute to the NPI's general program in drug abuse research  by complementing the Institute's biobehavioral perspective.    During the award period, the applicant also expects to grow professionally  as Associate Director of the UCLA Drug Abuse Research Center. in addition  to pursuing the aforementioned research. activities will include the career  development of new investigators from various disciplines and the mentoring  of graduate and undergraduate students in related fields. | https://reporter.nih.gov/project-details/2115970 | K02 | 1995 | 0.69 |
| DESCRIPTION: (Adapted from the applicant's abstract) This project  proposes a biosensor for the detection of DNA fragments specific to HIV.  The sensing element will be an interdigitated array (IDA) having single  stranded DNA immobilized between the array's digits. Immobilized DNA  strands will be complimentary to a conserved sequence in the HIV target  thus allowing selective detection. Electronics capable of measuring  impedance changes will be the detection of the hybridization of  complementary DNA strands. Impedance spectroscopy, i.e., multiple  frequency impedance measurements, will be used to completely characterize  the sensor's response. The resulting impedance spectrum will be treated  by Neural Network Analysis (NNA) to isolate the hybridized DNA response  from non-specifically adsorbed species and other matrix effects. The  technological innovations which will result from the project include a  sensitive and specific HIV DNA detection scheme which would be applicable  to any DNA assay. | https://reporter.nih.gov/project-details/2005782 | R41 | 1997 | 0.69 |

**Cluster 28:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
| The long term goal of our research is to understand the flow of  information from the genome to the phenotype of organisms. In this  proposal, we will attempt to use Bayesian networks and near-optimal  sequence alignments to represent protein secondary structures and motifs.  A Bayesian network describes the likelihood of amino acids at each  position in a motif as well as the dependence of amino acids in one  position on the amino acids at other position. Hence, Bayesian networks  can describe both the conservation of amino acids at single positions and  the conservation of correlations between two positions simultaneously.    Conserved amino acids result from evolutionary selection for a specific  amino acid or type of amino acid at one position in a protein structure.  These positions often have important functional or structural  requirements. Correlated changes between amino acids generally result from  side-chain side-chain interactions between pairs of amino acids in a  protein's structure. The types of correlations we have represented with  Bayesian networks include electrostatic charges, hydrophobicity, hydrogen-  bond donor and acceptor and inversely correlated packing volumes among  others. These Bayesian networks can be used to 1) discover side-chain  side--chain interactions within protei motifs and 2) to search sequence  databases for motifs showing both correlations and conserved amino acids.    Near-optimal alignments between two sequences can display regions that  have been more highly conserved or less highly conserved using the  information contained in only two sequences. The most highly conserved  region correspond to the most highly structured regions and the most  highly variable regions correspond to loops and coils and other  hypervariable regions. We propose to use near-optimal alignments to  display conserved secondary structures of proteins and hypervariable  regions. We will use secondary-structure specific amino acid substitution  matrices to provide specificity.    The goals of this proposal are to 1) build a database of Bayesian networks  that represent protein motifs, 2) test these networks for their ability to  detect motifs using test sets and crossvalidation methods, 3) compare  these networks with other methods for searching protein databases , 4)  build an integrated set of Bayesian networks to predict protein secondary  structure, 5) compare the prediction of protein secondary structure with  existing method 6) build a near-optimal sequence alignment workbench, and  7) predict structured and unstructured regions in proteins from near-  optimal alignments. | https://reporter.nih.gov/project-details/2519669 | R01 | 1997 | 0.91 |
| This research addresses one of the remaining challenges in speech science,  high quality speech simulation. We define speech simulation as a form of  speech synthesis in which the movement of air and tissue is under  experimental control, rather than the resulting acoustic signal. From the  early days of speech synthesis nearly a half century ago, the expectation  has always been that a better representation of the laws of physics of air  and tissue in motion would produce better synthesis. Although this  expectation still exists today, the payoff has been slow, primarily  because there are few data sets from which to build theoretical  generalizations. In this proposal, the principal investigator and his  colleagues draw upon experience gained with simulation of the phonatory  processes to include the entire vocal tract in sentence-level speech  production. The first phase will be to obtain naturalness in speech  quality that is comparable to formant synthesis by modeling a few specific  speakers from whom extensive data sets will be available. The second  phase will be to develop scaling and modification rules that will allow  the voice of a given speaker to be transformed into a different age,  gender, emotion, and voice quality. The transformation will also include  induced or corrected voice and speech disorders. The idea of voice  transformation (conversion) is not new, but the attempt to do it all in  the articulatory domain is relatively untried. The results will have  practical and theoretical impact on the development of assistive devices  for voice/speech impaired populations, for surgery performed on the larynx  and upper respiratory tract, and for speech training and rehabilitation. | https://reporter.nih.gov/project-details/2443629 | R01 | 1997 | 0.91 |
| DESCRIPTION: The principal investigator notes that the Aldol addition  reaction is one of the most important carbon-carbon bond forming  reactions in synthetic organic chemistry, that it is widely used in the  synthesis of drugs and other biologically active molecules and that the  reaction is now most often run in non-polar solvents, such as  tetrahydrofuran (THF), frequently with lithium salts of carbonyl  compounds (lithium enolates). He reports that the aggregation of these  compounds has been characterized by solution properties, NMR and by X-ray  crystallography and that such aggregates have been proposed as the  reactive species in the Aldol-type additions with other carbonyl  derivatives but there have been very few studies on the actual role of  such aggregates in reactions. He goes on to note that by a combination  of UV-vis spectroscopic and proton transfer equilibrium studies of some  lithium enolates, aggregation constants have been obtained even in  dilute solution and that spectra as a function of concentration are  analyzed by "Singular Value Decomposition" to determine the number of  different species present and to permit deconvolution to give the  spectrum of each component. For this purpose a glovebox-spectrometer  apparatus has been developed in which the sample compartment built into  the glovebox is connected with a spectrometer with fiber-optic cables.  It is noted that the apparatus permits spectroscopy of solutions  prepared and studied under the inert atmosphere of the glovebox. The  results thus far are said to suggest that the monomeric ion pairs might  play an important kinetic role in reactions.    It is proposed to extend such studies to additional enolates of interest  and to measure the reaction kinetics of their Aldol additions to  aldehydes and ketones and of Michael addition reactions with unsaturated  carbonyl compounds. It is indicated that the kinetic studies will show  the state of aggregation of the enolate reactive species and that  knowledge of the relative roles of ion pair monomers and aggregates will  lead to more complete reaction mechanisms and to the better  understanding required for sophisticated synthesis design. The principal  investigator notes that in particular, the roles of solvent addends,  such as lithium salts, hexamethylphosphoric triamide, and di- and  triamines will be studied under carefully controlled conditions to  determine the role of coordination of lithium in the stereochemistry of  the addition reactions.    It is also proposed to apply the same techniques of spectroscopic study,  proton transfer equilibria and reaction kinetics of Aldol and Michael  addition reactions to the dilithium salts of carboxylic acids and beta-  diketones. It is noted that these dilithium salts are also being used  increasingly in organic synthesis but that the reaction mechanisms are  virtually unstudied. It is reported that these salts are also aggregated  but nothing is known about the relative reactivities of monomers and  aggregates. The proposed studies are to provide unique information  about these reactions, which would be difficult to obtain in any other  way. It is suggested that subsequent extension to other salts of alkali  and alkaline earth metals, early and late transition metals and  lanthanides is also proposed since many of these salts have found use  in some stereospecific syntheses. | https://reporter.nih.gov/project-details/2608789 | R01 | 1998 | 0.91 |
| Automated screening of Pap smear slides is challenging due the high  processing and data transfer requirements placed on the processing  engine. These requirements can be significantly reduced by processing  the images at the image plane of the camera, and reading out only the  relevant data, which results in lower cost and higher performance  systems. Image sensors with smarts or computational capability at each  pixel can be advantageously used in the application to build extremely  compact and low-cost automated screening systems. Morphological  filtering algorithms have been shown to be effective at detecting object  size and shape, which are distinguishing features in diagnostic  microscopy. The goal of this research is to design, simulate and  fabricate a CMOS chip with morphological filtering circuits at each  pixel, which will allow detecting suspicious cells in a Pap Smear at  more than 1000 frames/second. In Phase I, Bosonics will (1) determine  the desired imager's technical capabilities for a compact microscope  mountable smart camera (2) design candidate morphological filtering  architectures and circuits; (3) simulate the morphological algorithms  with realistic circuits models; and (4) design and fabricate a 5x5  imager demonstration chip in CMOS.    PROPOSED COMMERCIAL APPLICATION:  The chips produced under this program will lower the cost of machine  vision systems by providing an integrated detector/processing function  as well as increasing performance by reducing the output bandwidth  requirements. The detector arrays developed under this program will have  wide application in automatic target recognition, machine vision for  automated manufacturing, medical diagnostic imaging, and remote sensing  and surveillance. | https://reporter.nih.gov/project-details/6012563 | R43 | 1999 | 0.91 |
| This application seeks funding for Phase II of a three-phase program of  research to identify, classify, and test nursing sensitive patient  outcomes and their indicators for use in standardized language  development, practice, research, and education. The purposes of the  research are to: 1) identify, label, validate, and classify nursing  sensitive patient outcomes and indicators, 2) evaluate the validity and  usefulness of the classification in clinical field testing, and 3) define  and test measurement procedures for the outcomes and indicators. The  classification is expected to contain patient outcomes, indicators, and  measurement activities at three to four levels of abstraction and to  identify those patient outcomes most influenced by nursing. The research  uses both inductive and deductive approaches. An inductive approach will  be used to extract outcomes, indicators, and measures from current  nursing literature and instruments. A combined inductive and deductive  approach will be used to label outcomes, specify indicators for the  outcomes, and group the outcomes in broad categories based on the Medical  Outcomes framework and categories identified by nurses. Delphi  techniques and surveys of random samples of masters prepared nurses will  be used to validate the outcomes and indicators prior to field testing  the outcomes and indicators in four sites, a tertiary care hospital, a  community hospital, a nursing home, and a community agency. Hierarchical  clustering techniques and nonmetric scaling analysis will be used to  develop the classification of nursing sensitive patient outcomes and a  survey of nurse experts will be used for initial validation of the  classification. | https://reporter.nih.gov/project-details/2257355 | R01 | 1995 | 0.66 |

**Cluster 29:**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Title | Awardee | Award Activity | Year | Sample Silhouette Score |
|  | https://reporter.nih.gov/project-details/2249310 | R43 | 1993 | 1 |
| This application is a request for a NIH Research Scientist Development  Award, Level II RSDA-II) to extend work supported under the RSDA-I  previously awarded to the applicant (DA00139). (The RSDA-I granted in 1989  was titled "Innovative Statistical Approaches to Drug Abuse Data"; this  application for the RSDA-II is titled "Drug Abuse: Epidemiology, Treatment  Processes, and Outcomes.") The RSDA-Il will continue to ensure financial  stability and release time from the pursuit of funding for actual research  work. The major focus for the applicant during this five-year award is  continuation of her professional work examining drug use epidemiology and  treatment interventions for problematic drug abuse. Examining the  implications of research findings for treatment strategies and developing  the necessary social policy changes to support the implementation of  improved treatment strategies is of further interest. The applicant will  continue her professional work applying innovative statistical  methodologies to drug abuse data. To this end, three convergent lines of  current research will be continued. The first examines drug use and  treatment utilization among subjects recruited through hospital emergency  rooms sexually transmitted disease clinics, and jails. The second project  is to improve the efficacy and efficiency of matching drug users' treatment  needs to services. The third involves examining and evaluating the process  of treatment service delivery with a special focus on the roles and  functions of drug treatment counselors.    The applicant's supporting institution is a research unit, the  Neuropsychiatric Institute (NPI), organized within the Department of  Psychiatry, School of Medicine, UCLA. Affiliated with the NPI is the UCLA  Drug Abuse Research Center, which has been conducting research in drug  abuse epidemiology natural history of narcotics addiction, treatment  evaluation, and social policy over the past 20 years. In this setting, the  applicant will conduct the proposed research and will receive additional  training in psychiatric aspects of drug abuse treatment and in the  implementation of treatment services. Furthermore. the applicant's  considerable psychosocial research knowledge and skills in drug abuse  issues will contribute to the NPI's general program in drug abuse research  by complementing the Institute's biobehavioral perspective.    During the award period, the applicant also expects to grow professionally  as Associate Director of the UCLA Drug Abuse Research Center. in addition  to pursuing the aforementioned research. activities will include the career  development of new investigators from various disciplines and the mentoring  of graduate and undergraduate students in related fields. | https://reporter.nih.gov/project-details/2458329 | K02 | 1997 | 1 |
| The goal of this project is to develop a computer algorithm that  accurately predicts the tertiary fold of a small protein given only  its amino acid sequence and secondary structure. It will use an  evolutionary optimization method with proprietary enhancements  invented by the PI. The specific aims of Phase I are to: (1)  develop a more natural representation of the problem; (2) implement  an energy function appropriate for this representation; (3) improve  the optimization algorithm; and (4) combine these into a simulator  that successfully predicts tertiary from secondary structure. In  Phase II this simulator will be extended to predict the tertiary  structures of larger proteins from their sequences and from sparse  experimental constraints or imperfect secondary structure  predictions. It will be validated using many proteins of different  sizes and fold classes and developed into software products for the  pharmaceutical and biotechnology industries. The long-term goal is  to help satisfy the rapidly growing demand for the accurate  prediction of protein structures from sequence information. This  demand arises not only from the exponential growth of the database  of sequences, but also from the need to understand the structure  and function of newly discovered gene products known to be involved  in human disease.    PROPOSED COMMERCIAL APPLICATIONS Commercial applications include  protein structure determination, structure-based drug design and  protein engineering in the pharmaceutical and biotechnology  industries and in basic academic research. | https://reporter.nih.gov/project-details/2423689 | R43 | 1997 | 1 |
| Non-small cell lung cancer (NSCLC) is the leading cause of cancer morality  in men and women in the United States, and the overall long-term survial is  less than 15%. Pathologic stage I makes up 25-35% of NSCLC cases and has  a good prognosis. However, cancer relapse and death rate in this subset is  35 to 50% by 5 years. Chemotherapy is beneficial for the treatment of  several localized solid tumors after resection and may prove to be useful  in the treatment of patients with stage I NSCLC. Thr purpose of this  project is to define tissue and serum tumor markers in patients with stage  I NSCLC which predict for early cancer recurrence. Pathologic stage I  NSCLC was chosen for study to eliminate the significant influence of  positive lymph nodes and distant metastases on survival.  Immunohistochemical staining will identify potential tissue tumor markers  and radioimmunoassay (RIA) or enzyme-link immunosorbent assay (ELISA) will  identify potential serum tumor markers.    Specific aim #1 will examine a set of twelve tissue tumor markers in a  retrospective cohort of 275 stage I NSCLC patients. Markers are  categorized by hypothetical method of action: molecular genetic markers  (Kras, erbB-1, erbB-2, rb, p53, bcl-2), markers of metastatic propensity  (angiogenesis factor viii), proliferation markers (K1-67) and markers of  cellular differentiation (Blood group A, H/LeV/LeB, NCAM, CD44). Results  will be used to develop a prediction rule for recurrence in stage I NSCLC  using Cox proportional hazards regression analysis and an artificial neural  network.    Specific aim #2 will examine a set of eight serum tumor markers in a  retrospective cohort of 250 patients with stage I to IV NSCLC. These  markers are categorized as molecular genetic markers (anti-p53), markers of  metastatic propensity (angiogenesis bFGF), somatamedins (growth factor IGF-  1) and markers of cellular differentiation (CEA, CA-125, CA 15-3, CYFRA21-  1, CD44). The purpose of this aim is to identify any correlations between  titers of serum markers and tumor histology, stage or mass. One hundred  patients in this cohort had a second serum collection after tumor  resection. This subgroup of serum will allow analyses of titters before  and after cyto-reduction. Significant correlates with tumor stage and mass  will be evaluated in a prospective cohort of patients with stage I NSCLC.    In specific aim #3, paraffin-embedded and fresh-frozen tumor tissue will be  collected from a prospective cohort of 330 patients with stage I NSCLC to  validate the prediction rule developed in specific aim #1. In these same  patients, serial serum specimens will be collected for a minimum of 2.0  years after resection (specific aim #4). The significant markers  identified in specific aim #2 will be analyzed in this cohort to describe  correlations with tumor recurrence. Tissue and serum markers identified by  the model can be used to select high risk patients for a prospective,  multi-institutional chemotherapy trial for stage I NSCLC. | https://reporter.nih.gov/project-details/2769931 | R29 | 1998 | 1 |
|  |  |  |  |  |