Narrow Research on the ECoS Paradigm - Interim Report 1

May 9, 2007

1 Introduction

Previous research activities on Artificial Intelligence (AI) and Machine Learning (ML) identified the ECoS paradigm (Evolving Connectionist Systems, introduced by [8]) as a possible key enabler for the proposed "morphing agent" approach. To be recalled shortly, the distributed decision taking process within the given unbalanced environment brings up several specific requirements which can be subsumed as the following optimization problem: balancing the workload and decision taking between weak mobile and powerful backend system nodes by providing maximum autonomous runtime of the mobile devices in contrary to being sensitive for alarm detection (patient emergency detecion).

Depending on the individual characteristics of the monitored object (i.e. the individual physical constitution of the patient), it is rather vague and impossible to specify how this optimization has to look like. In contrary of having multiple instances of learners on multiple nodes - which brings up the problem of having a superior learning intelligence managing the optimization of the distributed learners - a single learning instance moving between nodes is proposed (called "morphing" agent because the agent *morphs* from one shape (a node) into another). The major advantage of this approach is that the complicated task of synchronizing the decision taking of multiple autonomous learners can be replaced by a single decision taking learner, which "simply" moves over the network from the weak mobile nodes to the powerful backend ones depending on the status of the monitored object. Moving or "morphing" the decision taker from one node to another implicitly includes moving its knowledge and current status (its “mind”), which seems to be possible by utilizing the ECoS paradigm.

See [14] for further description of the proposed morphing agent approach.

To gather deeper knowledge about the ECoS principle, its implementations, strengths and weaknesses and research activities in other projects with similar requirements, a narrowed research tailored to the topic of evolving connectionist systems (ECoS) is performed. As online/real time/life-long machine learning is a rather young field, literature on that topic is not (yet) available in masses. Nevertheless a couple of interesting papers provide a good insight in several available implementations of the ECoS paradigm (e.g. EFuNN, DENFIS, SECoS), their application on real-world projects and alternative approaches to the ECoS paradigm proposed by other researchers.

2 Foundations / Implementations

The foundations of the ECoS paradigm of Nikola Kasabov is bound to the (meanwhile present) and future requirements for connectionist systems. These tasks usually require flexible learning and dynamically adaptive connectionist systems (COS, i.e. neural network structures) that have 'open' structures and are able to process both data and knowledge. Seven major requirements (that are addressed in the ECOS framework) are listed below (taken from [8]):

1. A COS should be able to learn quickly from large amount of data therefore using fast training, e.g. 'one-pass' training.

2. A COS should be able to adapt in a real time and in an on-line mode where new data is accommodated as it comes.

3. A COS should have an 'open' structure where new features (relevant to the task) can be introduced at a later stage of the system's operation, e.g., the system creates 'on the fly' new inputs, outputs, connections and nodes.

4. A COS should be able to accommodate in an incremental way everything that is, and that will become known about the problem, i.e. in a supervised (instructional), or in an unsupervised mode, using one modality or another, accommodating data, rules, text, image, etc.

5. A COS should be able to learn and improve through active interaction with other COSs and with the environment in a multi-modular, hierarchical fashion.

6. A COS should adequately represent space and time in their different scales, inner spatial representation, short- and long-term memory, age, forgetting, etc.

7. A COS should be able to analyse itself in terms of behaviour, global error and success; to explain what it has learned and what it 'knows' about the problem it is trained to solve; to make decision about its own improvement.

A good description of the fundamentals of ECoS and the representative EFuNN implementation can be found in [9]. It points out that a significant advantage of ECoS implementations like EFuNN is the local learning algorithm and also mentions that one of the problems is to find optimal values for the evolving parameters (sensitivity, error treshhold, learning rate). EFuNN investigated in more detail can be found in [10]. This paper provides good theoretical background for implementing ECoS/EFuNN yourself.

[11]compares two architectures of knowledge based neural networks: FuNN and EFuNN. It points out that EFuNNs are much faster than traditional FuNNs and therefore applicable for online learning.

A detailed description of DENFIS (dynamic evolving neuro-fuzzy inference system) is presented in [12]. DEFNIS is a competitor/successor to EFuNN and can achieve better performance under certain circumstances. DENFIS also outperforms the popular ANFIS (which is no on-line learner) [//TODO: reference ANFIS].

3 Applications / related research

[15]for instance built a predictive Evolving Fuzzy Neural Network (EFuNN) which predicts step t+1 based on its life-long experience. He found out that his EFuNN-T would perform significantly better, if it could deal with patterns in memory/history of previous prediction experience. This would require the integration of the so called "automaton" (a recurrent network structure) into the EFuNN.

[20]et al. show the ECoS paradigms / EFuNNs strengths being utilized as classifier with less training data available. They built an image classification system for pest damage recognition of apples based on the EFuNN structure. Advantages are that it works with only less training data and that rules can be extracted for further analysis. The authors found out that an important issue

for the performance of the EFuNN is the finding of a suitable feature selection method.

As system for analysis of gene expression data and its correlation to diseases was developed by Kasabov et. al in [7]. In this study one of the major problems were the large data sets and the importance of a timely response of the system. The approach performs clustering and filtering of input data to achieve a responsive system by maintaining accurate results. A hierarchy of EFuNNs was implemented: First, a "mother" EFuNN is trained on the large dataset offline. The mother is used to analyse features of interest. In a second step the rules are extracted out of the mother EFuNN. In a next step, a model is built based on the rules learned by the mother EF. Input variables of the extracted rules are ranked, the input data set is reduced to variables with a rank greater than a defined treshhold. After the model is built, a second EF is initialized with the model built before, the second EF is then deployed for production usage to the live system.

[2]proposes an adaptive speed recognition system based on ECoS. The basic requirment of this research is that the classifier is required to be able to identify new classes at runtime. SECoS was used here too. GHO states that one advante of local optimising learners like ECoS (SECoS) is that knowledge in nodes only represent a "patch" of input-output space, so additinoal output-classes can easily be added to an existing ECoS without the need to retrain the whole system as it would be rquired for traditional ANN systems. The expansion of the output classes is performed in a supervised way, i.e. is only performed if input vector is identified to belong to a new class.

[6]is dealing with multimodal information systems, i.e. builds a classifier for speech recognition that integration more than one modality of information (e.g. speech recognition with modality of visual data of the lips and sounds of spoken words). A general problem of multimodal system is synchronisation of modalities. Currently it seems to be out of scope of this research project.

[21]researched the usage of evolutionary computation in computer games to create interactive opponents. This research itself does not investigate the ECoS principle, but rather general GA-based AI approaches but some interesting literature references for futher reading are given

4 Problems of ECoS and Implementations

[3]'s work concentrates on a similar topic: reduce number of input variables with large dimensional data sets with accurate classification performance. He cites Ramsey & Schaefer 2002: "reducing features to significant ones improves inference and classification" which clearly emphasises my notion that feature selection is a critical fundament for the quality of the system. GOHs approach is to apply multiple feature selection methods in a hybrid way to reduce redundant variables.

A major problem of ECoS systems is the optimal parameter definition (e.g. learning rate, treshholds, etc.). The definition of optimal parameter values is always a trade-off e.g. sensitivity vs. classification performance, learning rate vs. forgetting, aggregation vs. destroying of knowledge, etc. [18] studied parameter optimisation of life-long learning ECoS systems by utilizing evolutionary approaches (GA). He state that a continous stream of input data may require the change of parameters in an on-line mode during the process. The ECoS implementation used for his investigatino was SECoS (Simple ECoS - a minimalistic ECoS implementation). The results of his investigation shoed that GA-optimised ECoS produced significantly smaller networks by keeping competitive to bigger un-optimised networks. The drawback of his research is that the GA-optimisation used is not suitable for real-time usage because of its high computational cost. Further investigations to overcome this problem is needed.

5 Alternative approaches

[4] proposed a fast algorithm for batch mode training of ANNs, the so called "extreme learning machine" ELM. A real-time/online variant of ELM is presented: OS-ELM. ELM is not tuning ANN parameters as ordinary ANN approaches do but rather randomly changes weights of the input neurons and biases of hidden neurons and then analytically determines output weights. ELM provides better generalization performance at extreme learning speed (at least regarding to HUA). The key to success is the calculation of the so called "pseudo-inverse" of the hidden layers output matrix. Fuzzy-ELM is a fuzzy variant of ELM. Compared to other approaches it was found that FUZ-ELM seems to be a valuable alternative to EFuNN. It has to be investigated if ELM fulfills the principles of ECoS which is the key enable for the morphing agent approach.

6 Conclusion and Further Reading

ECoS implementations are already available in a wide variety of applications for online classification and control. The major strengths of ECoS based systems are:

- fast, single-pass, online-learning
- ability to work with only minimal amounts of training data
- ability to add new output classes at runtime
- being robust to forgetting and over-training
- achieving accurate classification and generalisation results, in most cases outperforming traditional approaches

Popular implementations for the ECoS paradigm are EFuNN (Evolving Fuzzy Neural Network), SECoS (Simple ECoS) and DENFIS (Dynamic Evolving Neuro-Fuzzy Inference System). They all share a similar local learning algorithm which seems to be the key point to fulfilling the ECoS paradigms requirements. Although ECoS-based systems have the potential to achieve excellent results, there are two major design challenges strongly influencing the performance of the outcoming sytem:

- definition of the evolving parameters (learning rate, tresholds, etc.)
- feature selection out of the incoming data set

Already existing approaches to deal with these problems are for instance to pre-analyze the input data sets (if available) in offline mode and try to find out features of interest. Evolving parameters can be tuned during system operation utilizing GA-based approaches.

Although I found many papers regarding ECoS and its implementations, most of them directly include the original inventor/author of the ECoS paradigm Nikola Kasabov. Due to the promising results of the presented ECoS based systems it is hardly explainable why the ECoS paradigm or EFuNN as a popular implementation has not yet been adopted by more researchers within their projects. I would suggest that the terminology "ECoS" is not a commonly known word whereas my literature research regarding ECoS could only return those rather "Kasabov-specific" results. There is definitely a very active researcher base working on online learning systems as the ECoS paradigm describes, e.g. the previously mentioned "extreme learning machine" ELM by HUA??. I would suggest the common ground of ECoS-based and related research is not the term "ECoS" but rather the fact of implementing local learning strategies.

I will extend this literature research phase to redo search and eval for papers with the more general "local learning" keywords instead of explicitly looking for "ECoS", the aim is to find related work of other researchers NOT being directly involved in ECoS projects.

- //TODO list papers of interest to be read now

7 Intermezzo - Smart Clothing

I read about researchers working on "smart clothes" which are capable of measuring vital values e.g. blood sugar rate. It has to mentioned in the thesis that this development towards several intelligent devices and clothes worn by humans does definitely NOT make this research irrelevant as external people could think. In fact the intelligent or smart clothes are acting as "datasource" for the intelligent patient monitoring algorithm as suggested. The only thing that is different is, that all sensor data is not delivered by a single belt worn device as initially assumed, but rather has to be collected by the intelligent monitoring device from built-in sensors and additional "external" sensors - i.e. smart clothes, watches, belts, etc. The interconnection and communication of those objects clearly opens up an additional field of research, "body area network" is an adequate keyword for this topic I assume.

8 Further reading of recent publications on ECoS, it's implementations, use-cases and alternative approaches

8.1 [WAT06b][19]

- -comparison of two ANN methods for prediction task (predicting the risk of insect pest species establishment in regions)
- -compares: kohonen self-organising-map (SOM) and simple ecos SECoS

- -both types perform spatial clustering of training examples during learning
- -the advantage of SOM: is to visualize the clusters that are formed
- -the disadvantages are that training is extremely slow and the number of clusters is fixed
- -clear advantage of secos: number of clusters is determined automatically during training and the fast speed with which training occurs

8.2 [KAS05][5]

- -incremental learning in autonomous systems-ecos for online image and speech recognition
- -a new type of ecos impl is presented: evolving growing cluster classifier (EGCC)
- -the implementation of the concept if growing in egcc is different from that of other growing neural networks, such as growing neural gas (gng) and growing cell structures (gcs) - egcc has simple training and test procedures that require only two-pass training and there is no need for a parameter setting
- -principle component analysis (pca) is a typical method for feature extraction
- -original pca is not suited to incremental learning purposes, incremental principle component analysis(ipca) is used
- -due to the requirement of two-pass learning of the egcc it seems to be not applicable for our research project

8.3 [LIA06][13]

- -online sequential learning algorithm OS-ELM is presented
- -OS-ELM is faster than other sequential algorithms and produces better generalization performance
- -although OS-ELM is NOT compared to ECoS (e.g . EFUNN,DENFIS,SECoS) in this paper

- -OS-ELM input weights and biases of neurons are randomly generated and based on this the output weigthsts are analytically determined.
- -OS-ELM only requires the number of hidden nodes to be specified, no further parameters have to be tuned
- -OS-ELM is compared to SGBP,RAN,RANEKF,MRAN,GAP-RBF,GGAP-RBF
- -apart from selecting th enumber of hidden nodes, no other control parameter has to be chosen

8.4 [ANG06][1]

- -a "rough set based pseudo-outer product (RSPOP)" stock trading decision model is presented
- -benchmarks are performed against several other approaches in stock trading prediction/decisino taking, including DENFIS (ECoS impl.)
- -increasingly applications based on AI have been applied to financial forecasting systems
- -although neural networks possess the properties required for technical finance forecasting, they cannot be used to explain the causal relationship between input and output variables because of their black-box nature
- -the main strength of NFS (neuro-fuzzy systems) is that they are universal approximations with the ability to solicit interpretable if-then rules.
- -the strength of feuro-fuzzy systems involves two contradictions: interpretability vs. accuracy
- -linguistic fuzzy modelling is focused on interpretability, mainly the mamdani model
- -precise fuzzy modelling is focused on accuracy, mainly the takagi-sugeno-kang (TSK) model
- -recently, rough set methods have been shown to significantly reduce pattern dimensionality and proved to be a viable data mining technique

- -the use of rspop algorithm reduced computational complexity, improved interpretability by identifiying significantly fewer fuzzy rules
- -the rspop algorithm consists of three parts: rule identification: identifies only the most influential rule instead of all possible rules (ANN: compare with M-of-N vs. 1-of-N rule activation stuff in ECOS!!!!)
- -attribute reduction (step 2) performs feature selection to perform reductino of redundant attributes using the concept of knowledge reduction of rough set theory
- -step 3: rule reduction performs partial feature selection by extending the reduction rules without redundant attributes
- -this approach was compared to: feed-forward neural networks with BP(FFNN-BP)
- -radial basis function networks RBFN
- -ANFIS
- -EFuNN
- -DENFIS
- -EFuNN and RSPOP-CRI yielded the best recall and predictive results respectively
- -EFuNN demonstrated an "over-trained" phenomenen at prediction results
- -RSPOP_CRI and DENFIS yielded recall and prediction results that are superior to the Random Walk Model
- -DENFIS yielded a slightly LOWER MSE and a slightly higher number of rules than RSPOP-CRI
- -denfis is based on TSK whereas POPFNN-CRI is based on mamdani which explains the differences
- -overall impression of this comparison is that DEFNIS performs slightly equal compared to the presented rough-set based method

- -the authors also conclude with "DEFNIS and a novel rough set based neuro-fuzzy system called RSPOP-FNN yielded superior predictive performance than the well-established random walk model." (and others)

8.5 [ZAN05][22]

- -design of experiments in neuro-fuzzy systems
- -the adaptability and instability of the problem demands constant system re-configuration, so that they are used in situations that the operational points of the problem comes in different ways
- -determination of network parameters is a hard designer task
- -a statistical analysis if accomplished to verify the interactions and interrelations between variables in the design of neuro-fuzzy systems and to verify the most relevant factors in the design of these systems. the method used to perform this analysis was the "Design Of Experiments" (DOE).
- -outcome: the most relevant parameters for EFuNN are membership funtion shape and number
- -in automatic optimization techniques the information can be mapped in the cost functions

8.6 [WAT00][17]

- -SECoS and experiments on isolated phoneme recognition
- -at first glance, the ecos paradigm may seem very similar to the Resource-allocation Network proposed by Platt [11] (-> cite).
- -closer inspection shows some fundamental differences,e.g.
- -ran is a more complex system as it has more parameters to adjust and requires more complex calculations
- -ran is not a one-pass learning algorithm
- -above 2 differences are critical success factors for our research project(less computing intensive, one-pass online learning), so RAN seems to be no alternative option to ECoS based systems

- -aggregation of hidden layers in secos succeeded in significantly reducing the size of the hidden layer and somewhat improving the generalisation accuracy, but severly reduces the accuracy over previously seen data
- -the wide range performance between phonemes strongy suggests that the secos model is sensitive to both training and aggregation parameters. the suggested future work therefore focuses on automatically and dynamically adapting both the training and aggregation parameters

8.7 [WAT06a][16]

- - nominal-scale evolving connectionist systems
- - the evolving connectionist system algorithm is extended to represent and learn nominal scale data
- - normally, nominal scale data cannot be used as input data for ANN
- - the usual approach is to represent nominal-scale variables orthogonally,each possible class of each variable is assigned its own input neuron
- - this is not feasible for open-end data streams in online-learning applicatios, e.g.
- - expansion of dimensionality of the input space: each value of each variable adds a dimension to the input space, this results in a "curse of dimensionality" for even small nb. of variables
- - independence of inputs: orthogonal reprsentation ignores the restriction that if a variable is on eclass, it cannot be another class
- - changing the number of classes: as each class is represented by an independent neuron it is not possible to add new classes dynamically (at runtime)
- - the author developed a variant of theECoS pardigm which can deal with nominal scale data, NECoS
- - the outcome: the developed NEcos works and performs well on classification problems, but does not perform well on function approximation problems

8.8 [VIL06]

- -an adaptive local learning based methodology for voltage regulation
- in distribution networks with dispersed generation is presented
- -the integration of dispersed generation systems into running
- distribution systems requires advanced intelliged voltage regulation
- -in this approach, the overall problem is treated as a constrained
- multiobjective optimization problem ans solved by standard nonlinear
- programming techniques or more advanced solution technologies e.g.
- fuzzy-based evolutionary algorithms (==> similarities to our research
- project)
- -a local learning algorithm is used, the so called "lazy learning"
- algorithm (LL)
- -a database is integrated that stores the most relevant historical
- solutions of the voltage regulation problem

9 Raw notes on papers

9.1 [WANXX] [15]

- - EFuNN-T is a predictive EFuNN
- - it performs "one-step-ahead" prediction (predicts next input based on current one)
- - prediction networks are required to modify their structure and params continuously during the process
- - EFuNN-T = EFuNN + pruning + rule extraction + next step prediction
- - how is the prediction performed? seems that it is just the output of the regular EFuNN?

- - regarding to WAN, EFuNN-T would perform significantly better if it could deal with patterns in memory/history of previous prediction experience. this would require to include a so called "automaton" into the EFuNN (recurrent network structure).
- - works in these steops: process all input vectors, update EFuNN network structure, feed "signals" (not nearer defined in his paper) into the network and perform the prediction (ie. input= x(t) output= x(t+1))

9.2 [WOO04] [20]

- - common image classification systems depend on enough available training data to deliver acceptable results
- - a new method of using an evolving, adaptive connectionist based system (EFuNN) is proposed, 2 major advantages:
- - works with less training data and
- - provides rules that can be extracted for further analysis
- - proposed system: "pest damage" system for new zealands apple growers
- - detects types of pest based on images of the fruit and leafs
- - extracts rules (derives rules usable for the apple grower)
- - there are not enough sample images available to train classical image-classification-systems
- - different tact is needed
- - interesting parts of the image are "cropped" and used as training sample for the EFuNN
- - i.e. there are made n training examples out of a single picture file
- - challenges: good feature extraction method has to be found, find a way how to process full sized images, select an appropriate classifier
- - EFuNN was chosen as classifier for its already known advantages
- - WOO and KAS performed a former image-classification comparation between EFuNN and SVM where EFuNN (54%) outperformed the SVM(42%)

- - previous study with another feature extraction and image processing method than the current one
- - comparation in this paper: k-means classifier(42,7%), MLP (multi-pass trained ANN) (21,91%), SVM (57,23%) compared to EFuNN 75% recognition rate!
- - obviously an important issue for the performance of the EFuNN is the finding of a suitable feature selection method, i.e. input vectors intothe EFuNN have to be representative about the processed data

9.3 [KAS03x] [7]

- analysation of gene expression data in correlation to diseases
- performs a clustering and filtering approach of input data
- timely response is important
- as important as accurate results
- many methods are slow on those large input data space
- do feature selection to reduce input space -> ECOS -> EFuNN:
- train a "mother" EF continouosly (accomodates all data)
- extract relevant features from mother (rule extraction)
- create a model based on features and output classes (EF2)
- steps:
- train mother EF on all data
- extract rules
- rank important input vars within rules (p1 5.2 modelling)
- reduce data set to vars with rank > treshold
- train second EF2
- deploy EF2 to information system

9.4 [GOH04] [3]

- base is the feature selection method used in KAS03x (EFuNN, ECF - evolving classification function)
- perhaps this is an enhanced algorithm as dated with 2004
- aim is to reduce input variables (features) with large dimensional data sets with accurate classification performance
- reducing features to significant ones improves inference and classification (Ramsey & Schaefer 2002)
- multiple feature selection methods are applied in a hybrid way to reduce redundant variables
- then ECF is applied to build a classifier with the selected remaining variables
- ECF: ECoS for classification tasks, 4 layers
- study had a very positive outcome:
- hybrid feature selection methods + ecos based classification achieves higher classification rates than previous approaches
- avg. rate: 91%

9.5 [WAT03?] [18]

- for life-long learning ECoS systems parameters are optimised using evolutionary approaches:
- one major problem of ecos is the optimal definition of parameter values
- continous stream of input data may require change of parameters in an on-line mode during the process
- ecos impl used: SECoS - simple ecos (as a simpler EFuNN)
- minimalis ECoS implementation, consists of 3 layers
- need of parameter optimization, e.g. sensitivity vs. classification/generalisation performance; learning rate vs. forgetting; aggregation vs. destroying of knowledge

- the result of ECoS parameter tuning through GA:
- produces significantly smaller networks
- competitive accuracy to bigger, un-opimised networks
- GA optimisation in this study is not really suited for on-line manner because of its high computational cost and the requirements of a training and a test data set
- further investigations to overcome the real-time inhibiting problems are to be performed by the authors

9.6 [GHO03] [2]

- adaptive speech recognition with ecos
- alternatives to ECoS: RAN[8], NN-MLP[12], cascade-correlation learning[1]
- problem of all of them: able to modify parts of their internal structures but not able to identify and learn new classes online, during operation of the system
- study used SECoS implementation
- "activation propagation" ECoS vs. traditional ANN: ECOS=OneOfN,ManyOfN; ANN=each node activation is propagated further
- OneOfN: only the winning one (most activated one)
- ManyOfN: only those with A > treshold
- SECoS = simple version of EFuNN
- it does NOT fuzzify the inputs for the sake of performance
- 3-layers input -> evolving layer -> output layer (each node in the output layer represents on class)
- aggregation merges nodes which are spatially close to each other
- aggregation controls the size of the evolving layer

- advantage of local optimised learners like ECoS : knowledge in nodes represents only a "patch" of input-output space, so additional output-classes can easily be added to an existing ECoS without the need to retrain the network again (with all data) as required for traditional connectionist systems / ANNs
- output expansion (add new classes) is supervised, i.e. is only performed if input vector is specified to belong to a new class
- results showed:
- -> high memorization of training data
- -> high generalisation
- -> robust to forgetting

9.7 [KAS99] [9]

- ECOs:
- good “base” paper with good description of fundamental ideas and EFuNN implementation
- good presentation of the basic idea of ECOS and the foundations
- presentation and detailed explanation of EFunn as an implementation of ECOS
- conlus:
- significant advantage is local learning algorithm (fast, one-pass learning)
- difficulty: finding of optimal vlaues for the evolving parameters (sensitivity, error tresh, learning rate(s),
- forgetting rate, pruning, etc)
- possible solution is to regularly apply GAs

9.8 [KAS01b] [10]

- foundations of EFuNN implementation (algorithms, formulas, other sources of similar approaches)
- -> theoretical background of ECOS / EFuNN implementation

9.9 [YAN05] [21]

- AI in Computer Games: Generating Interesting Interactive Opponents by the use of Evolutionary Computation
- - real-time learning should be used to create adaptive computer opponents to increase game satisfaction of human players
- - identified real-time learning challenges: performance, fast adaptation
- - evolutionary learning methods are used, "co-evolution techniques" (combined offline and online learning)
- - evolving artifical neural networks EANNs are presented, possible alternative to ECOS/EFuNN?
- check:
- X. YAO, Y. LIU: towards designing artificial neural networks by evolution (appl. math. computation) 1996
- x. yao "evolutionary artificial neural networks" in a kent j.g. wililams c.m. hall (ed's) encyclopedia of computer science
- and technology vol 33 pages 137-170;
- x.yao "evolving artificial neural networks" proceedings of the IEEE vol. 87 1999
- supervised learning of ANN requires appropriate set of trainnig data which is sometimes either not available (as in online patient monitoring)
- because of evol. approaches effectiveness in such spaces, research on ANN training by means of artificial evolution has been very active (yao 1999)
- how is "online-learning" performed in YAN05's experiments:
- x) initial generatino of opponents is generated by offline-learning (GA)
- x) within "real time" of the game, the computer bots fitness is measured by means of quality/fitness of interaction with the human player
- x) then the fittest bot is cloned and replaces the least-fit bot
- -> i.e. the population is not replaced by a whole new generation, evolving the population is performed "step-by-step" in small iterations 1 by 1 during gameplay.

9.10 [HUA05] [4]

- Huang, et al [1,2,3] proposed a fast algorithm for batch-mode training of ANNs, so called "extreme learning machine" ELM.
- ELM is not "tuning" the ANNs learning parameters as classical algorithms do, but rather randomly changes weights of the input neurons and biases of hidden neurons and then analytically determines output weigths
- A modification of the ELM for on-line learning on sequential data OS-ELM is presented
- batch ELM algo in short:
- 1) assign arbitrary input weight and bias
- 2) calculate hidden layer output matrix
- 3) estimate output weight
- ELM provides better generalization performance at extreme learning speed. The key is the calculation of the so called "pseudo inverse" of the hidden layer output matrix
- sequential ELM is achieved through integratino of the "recursive least squares" algorithm, RLS
- OS-ELM uses a "boosting" phase which shall boost output performance by doing a quick and fast batch-training at bootstrapping time.
- Further, a fuzzy veriant of OS-ELM is presented: Fuzzy-ELM
- Performance:
- OS-ELM and FUZ-ELM are compared to with other popular algorithms e.g. MRAN,BP,GAP-RBF / ESOM, EFuNN, DENFIS
- Fuzzy-ELM basically provides better generalization by smaller network sizes than ESOM (evolving self-org. maps) and EFuNN and DENFIS.
- DENFIS is comparable performance as zero-fuzzy-ELM.
- First-Fuzzy-ELM has smaller network size and a bit better performance as DEFNIS.

- difference of zero-fuzzy-ELM and first-fuzzy-ELM? (First-Order TSK, Zero-Order TSK)
- conclusion:
- FUZ-ELM seem to be a valuable alternative to EFuNN. It is to be investigated whether FUZ-ELM/OS-ELM fulfull the requirements of the ECoS
- paradigm which is the key-enabler of the morphing agent approach.

9.11 [KAS00] [6]

- multimodal based information systems achieve better results than unimodal ones, e.g. speech recognition supported by modalitiy of visual data of the lips or visual object tracking in images supported by modality of the objects sounds.
- a general problem of multimodal information systems is synchronization of the modalities, i.e. data streams.
- the proposed framework integrates sub-related systems with one super-related system, that takes the inputs of the sub-systems as input for the decision making (i.e. a "preprocessor" system for each modality":
- Audio –| | –> descision making connectionist system | | Visual -|

9.12 [WAT04] - summary

- only abstract + bibliography available
- references of interest:
- Hamker, F.H. (2001). Life-long learning cell structures-continuously learning without catastrophic inference, Neural Networks, 14:551-573
- Kasabov, N.; Song, Q.: 2000, Dynamic evolving neuro-fuzzy inference system (DENFIS): On-line learning and application for time-series prediction. In Proceedings of the 6th intl. Conference on Soft COmputing, October 1-4, 2000, Fuzzy Logic Systems Institute, Lizuka, Japan, pages 696-702
- Kasabov, N.; Song, Q.: 2002, DEFNIS: Dynamic evolving neural-fuzzy inference systems. IEEE Transactions on Fuzzy Systems, 10(2):144-154

- Rantala, J.; Koivito, H. (2002): Optimised subtractive clustering for neuro-fuzzy models. In Proceedings of the 3rd international Conference on Fuzzy Sets & Fuzzy Systems (FSFS'02) Interlaken, Switzerland, February 11-15

- Romero, E., Alquezar, R. (2002): A new incremental method for function approximation using feed-forward neural networks. In Proc. of the intl. joint Conf. on Neural Networks (IJCCN) 2002, pages 1968-1973

- Wang, X. (2001): "on-line" time series prediction system EFunn-T. In Proc. of the fifth biannual Conference on Artificial Neural Networks and Expert Systems (ANNES2001), pages 82-86

- Watts, M.; Kasabov, N. (2000): Simple Evolving connectionist systems and experiments in isolated phoneme recognition. In Proc. of the first IEEE conf. on evolutionary computation and neural networks, San Antonio, May2000, pages 232-239. IEEE Press.

9.13 [KAS01] [11]

- detailed description of how efunn works - compares 2 architectures of KBNN (knowledge based NNs) : FuNN and EFuNN - efunns are faster (applicable for on-line learning) without compromising accuracy - good description of the EFuNN algorithm

- (already summarized somewhere in earlier notes)

9.14 [KAS02] [12]

- detailed description of DENFIS "dynamic evolving neural-fuzzy inference system" - starts with description of ecos and efunn - many refs to further literature - describes ECM "evolving clustering method" - denfis uses takagi-sugeno type inference - denfis uses m "highly activated rules" for dynamic (runtime) inference - defnis performs better than ANFIS in case of offline learning, better than several popular approaches in on-line learning - including EFuNN

- (already summarized somewhere in earlier notes)

9.15 [KAS98] [8]

- the "base" idea for the ECOS principle (already summarized somewhere in earlier notes)

References

[1] C. Ang, K.K.; Quek. Stock trading using rspop: A novel rough set-based neuro-fuzzy approach. *IEEE Transactions on Neural Networks*, 17(5):1301–1315, 2006.

[2] M.; Kasabov-N. Ghobakhlou, A.; Watts. Adaptive speech recognition with evolving connectionist systems. *Elsevier Information Sciences*, (156):71–83, 2003.

[3] Q.; Kasabov-N. Goh, L.; Song. A novel feature selection method to improve classifcation of gene expression data. *2nd Asia-Pacific Bioinformatics Conference (APBC2004)*, pages 161–166, 2004.

[4] N.Y.; Rong-H.J.; Saratchandran P.; Sundararajan N. Huang, G.B.; Liang. On-line sequential extreme learning machine. *the IASTED International Conference on Computational Intelligence (CI 2005), Calgary, Canada, July 4-6, 2005*, 2005.

[5] D.; Pang-P.S. Kasabov, N.; Zhang. Incremental learning in autonomous systems: Evolving connectionist systems for on-line image and speech recognition. *2005 IEEE Workshop on Advanced Robotics and Its Social Impacts*, pages 120–125, 2005.

[6] E.; Herik-van den J. Kasabov, N.; Postma. Avis: a connectionist-based framework for integrated auditory and visual information processing. *Elsevier Information Sciences*, 123:127–148, 2000.

[7] M.; Lane-T. Kasabov, N.; Middlemiss. A generic connectionist -based method for on-line feature selection and modelling with a case study of gene expression. *Knowledge Engineering and Discovery Research Institute*, 2003? School of Information Technology, Auckland University of Technology.

[8] N. Kasabov. Ecos: Evolving connectionist systems and the eco learning paradigm. In *Proceedings of ICONIP'98, Kitakyushu*. Department of Information Science, University of Otago, October 1998.

[9] N. Kasabov. Evolving connectionist systems for on-line, knowledge-based learning: Principles and applications. *The Information Science Discussion Paper Series*, (99/02), March 1999.

[10] N. Kasabov. Evolving fuzzy neural networks for supervised/unsupervised on-line, knowledge based learning. *IEEE Transactions of Systems, Man and Cybernetics*, 31(6), 2001.

[11] N.K. Kasabov. On-line learning, reasoning, rule extraction and aggregation in locally optimized evolving fuzzy neural networks. *Elsevier Neurocomputing*, 41:2545, 2001.

[12] Q. Kasabov, N.K.; Song. Denfis: dynamic evolving neural-fuzzy inference system and itsapplication for time-series prediction. *IEEE Transactions on Fuzzy Systems*, 10(2):144–154, April 2002.

[13] G.-B.; Saratchandran P.; Sundararajan N. Liang, N.-Y.; Huang. A fast and accurate online sequential learning algorithm for feedforward neural networks. *IEEE Transactions on Neural Networks*, 17(6):1411–1423, 2006.

[14] H.J. Prueller. Research notes: Model development (specific ai applicability research). 12 2006.

[15] Xin Wang. Online time-series prediction system: Efunn-t. *Department of Information Science, University of Otago, Dunedin, New Zealand*, 2001(?). xinw@infoscience.otago.ac.nz.

[16] M.J. Watts. Nominal-scale evolving connectionist systems. *Neural Networks, 2006. IJCNN '06. International Joint Conference on Neural Networks*, pages 4057–4061, 2006.

[17] N. Watts, M.; Kasabov. Simple evolving connectionist systems and experiments on isolated phoneme recognition. *Combinations of Evolutionary Computation and Neural Networks, 2000 IEEE Symposium*, pages 232–239, 2000.

[18] N. Watts, M.; Kasabov. Evolutionary optimisation of evolving connectionist systems. 2003(?).

[19] S.P. Watts, M. J.; Worner. Comparison of a self organising map and simple evolving connectionist system for predicting insect pest establishment. *International Journal of Information Technology*, 12(6):35–42, 2006.

[20] D.; Benwell-G.L. Woodford, B.J.; Deng. A wavelet-based neuro-fuzzy system for data mining small image sets. *ACSW Frontiers*, 32:139–143, 2004.

[21] G.N. Yannakakis. *AI in Computer Games: Generating Interesting Interactive Opponents by the use of Evolutionary Computation*. PhD thesis, The University of Edinburgh, 2005.

[22] F.L.; Ludermir-T. B. Zanchettin, C.; Minku. Design of experiments in neuro-fuzzy systems. *Hybrid Intelligent Systems, 2005. Fifth International Conference on Hybrid Intelligent Systems*, 2005.