Cognitive and Metacognitive Educational Systems:
Papers from the AAAI Fall Symposium (FS-09-02)

Capturing Knowledge in Real-Time ICT Systems to Boost Business Performance

Nadia Brancati, Giovanni Mappa

ANOVA Lab (Independent Researchers) - Centro Direzionale isola G1 – scala C, Napoli, Italy
Tel. 0817502535
n.brancati@anova.it, g.mappa@anova.it

Abstract

In this work an AI/ICT Platform is presented, to develop cognitive networks to cope with a management of a great availability of data and a necessity to dispose of prompt right information, extracted by data. In fact, the better strategic decision arise by a prompt availability of target and effective information.

A cognitive network, and in particular an intelligent grid, helps to reach this goal. This intelligent grid allows to integrate many data source to drive analytics which transform data into useful information to support advanced operational control and strategic decision making.

To realize an intelligent grid, it is necessary, firstly, capturing Knowledge, transforming data in information and introducing the knowledge in ICT framework and in Real-Time Systems. This is the right way to have a set of target and suitable information by using to take a correct decision, especially in real-time problem.

So, in this work XBASE Cognitive Mapping Tool is presented. This tool allows to develop an intelligent grid, to support and "automate" strategic decision and so, to solve, also in real-time, every kind of problems. In particular, an application of this tool is presented, in monitoring of wastewater, the "BATTLE" Project.

Keywords: cognitive network, intelligent grid, knowledge base, capturing knowledge, knowledge engineering, real-time system, expert system, environment monitoring.

1. Introduction

At present, there is a great availability of data. So, it is of primary importance that this mine of data is transformed in information and so, classified and synthesized. These information form the Knowledge, so the valorization of the target information can help to solve problems of different nature. To find a solution, above all for real-time problem, it needs to synthesize the Knowledge and introduce it in ICT and Real-Time Systems.

In fact, the Knowledge is the most important key of competition in the world wide market (Cheung and Liu, 2005 – Hia and Zhong, 2006). So, the keystone is to formalize and synthesize the Knowledge, using the Knowledge engineering.

Knowledge engineering helps to capture the Knowledge and insert it in ICT and in Real-Time systems. Knowledge engineering is a field within Artificial Intelligence that refers to the building, maintaining and development of Knowledge-Based Systems. Such systems are computer programs that contain a large amounts of Knowledge, rules and reasoning mechanisms to provide solutions to real world's problems. A major form of Knowledge-based system is an Expert System, one designed to emulate the reasoning processes of an expert practitioner, i.e. someone having preformed in a professional role for very many years.

In fact, Knowledge engineering also concerns mathematical logic, where the Knowledge is produced by socio-cognitive aggregates (mainly humans) and is structured, according to understanding of how human reasoning works.

In particular, Knowledge Engineering (KE) is related with the ability to scaling down a prefixed complex and implicit Knowledge Base (KB) toward a delivering explicit target information, in order that

$$\lim_{KE\to\infty}(KB) \approx 0 \tag{1}$$

In math words, given a prefixed problem solving domain, one may say that "the limit of a Knowledge Base for the Knowledge engineering process going on to infinity, reduce itself more and more, till zero". This assertion expresses the concept of Knowledge's synthesis.

Generally, this is possible by using a System Approach and Conceptual Tools of Artificial Intelligence, like Data Mining, Knowledge Extraction, etc.

An example of the assertion (1) could be what happened in the last year, in the mobile phone's area. In fact, the size of mobile phone has become more and more small. This has happened because the Knowledge has been synthesized,

and so the circuits have become more and more small, and so also the mobile phone.
A cognitive network, using the Knowledge engineering, is an incremental, innovative and essential solution to point out relevant and target information. An intelligent grid is a particular cognitive network, and it allows to transform data in information an to emphasize and synthesize the target information, in order to realize real-time system to solve problems, for example problems about industrial processes. This kind of network allows to "automate" the decisional processes, simplifying the search of problem solution.
XBASE Cognitive Mapping Tool, presented in this work, allows to develop an intelligent grid. It is introduced as a tool for expert systems development, having a role potentially equal to that the data flow diagrams widely used in information systems development. The cognitive map is used to provide feedback to domain expert, merge the Knowledge of multiple experts, synthesize this Knowledge and provide a graphic representation from which the final rule-based is formed.
XBASE Tool uses a cognitive network, with expert neurons.
In the second paragraph the functionality of a cognitive network are introduced. In the paragraph 3, XBASE Tool, and its software platform are described in details. In the forth paragraph a case study of XBASE is described. In the last paragraph some conclusions are drawn.

2. Cognitive Network

Generally, the artificial neural networks try to emulate some cognitive human processes. A neural network is made up of artificial neurons, called also nodes, and interconnections among nodes, and a numeric weight is associated to each interconnection. In every neural network, a single activation function (a sigmoid, a staircase function, etc.) exists (figure 1(a)). This function represents an elementary information, and according to the weights of input nodes, it returns the weights of output nodes. An artificial neural network is deterministic.
The real world is non deterministic, so it needs to find a non deterministic way to solve the real problem. A cognitive network uses a non deterministic approach. An important subclass of methods for reasoning with uncertainty is called "fuzzy logic" and the systems that use them are known as "fuzzy systems". In the most general sense, fuzzy logic is a multi-valued logic to express different degrees of certainty or uncertainty of assertions, so it can be used to solve non deterministic problem (Mappa et al. 1997).
In particular, in a cognitive network (or fuzzy neural network) there is a reticular structure, so as a classic neural network, but there is a set of activation functions, rather a single activation function. Every elementary node represents a Knowledge base. The interconnections among these neurons allow to fuse different cognitive processes. In these cognitive network, every node represents an homogeneous domain of Knowledge. In fact, the landmark of a cognitive network is a "Knowledge's node": an independent node, where primary data merge into. These data are transformed, firstly, in information, later in an homogeneous Knowledge Base, specific for the source area of data.
So, starting to elementary concepts, the interconnections create, dynamically, new more evolved levels of Knowledge's synthesis. The final level will produce the rules and actions of the expert system. The final complex output is produced as sum of simple elementary input. At each level, the Knowledge is synthesized, so at final level, one arrives to observe (1).
This process allows to solve, completely in automatic way, every kind of problems (Mappa et al. 1997).
A cognitive network, and in particular an intelligent grid, offers the follow advantages:

• it isn't composed by a single activation function, but by a matrix of activation functions;

• it is a "clear" system, and not a "black box", since one knows the application domain, and the Knowledge is introduced in the node of the network;

• it is very efficiency, since different activation functions exist;

• every expert neuron has a correlation with an homogeneous subdomain;

• every node allows to transform an input in a definitive or partial result, called meta-information;

• it allows to represent heterogeneous and extended Knowledge's domain;

• it operates with a various set of input, both qualitative than quantitative;

• at each step, it allows to obtain a grater level of Knowledge synthesis.

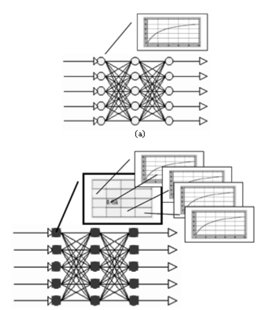

Figure 1: (a) artificial neural network; (b) artificial fuzzy neural network.

In figure 1(a) is visualized the functionality of a network with a single activation function. In figure 1(b) is visualized the functionality of a fuzzy neural network, in this case XBASE, with different activation functions.
In an intelligent grid, it is possible to specialize a first neuron's level, about a specific homogeneous Knowledge domain. It is possible to imagine an intermediate level, made of correlation matrices, which allow to synthesize the

information. These meta-information will be transfer to the final level, which will produce the final output, and so the answers and solutions of the problem.

To work with a cognitive network, and to automate completely this network, it needs to transfer the Knowledge from humans to information systems, and in particular, to every node of neural network. A first solution, to solve this problem, is the acquisition, the representation and the formalization of the Knowledge.

Any system of Knowledge, complex or not, can be represented and formalized in the follow way:

• collection of the information about every specific domain and case history of the problem;

• definition of the diagnostic problem and user requirements;

• definition of principal interconnections among elementary information;

• formalization of the proposed model by logic algorithm;

• simulation of the functionality of the proposed model and evaluation of the agreement between the model and a real problem;

• use of the model and analysis of the strategies that the model proposes to solve the real problem.

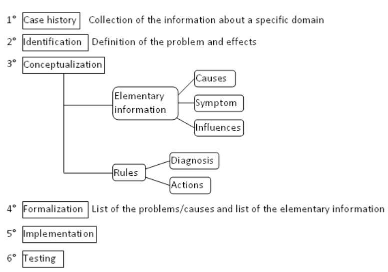

Figure 2: description of the formalization's process

In figure 2, a diagram describing the steps of formalization's process is visualized.

3. XBASE Tool

3.1 Description of the proposed method

The proposed method is based on "FuzzyMQC", Fuzzy Matrix of Certainty's Quantification. The Knowledge domain is split in homogeneous Knowledge subdomains. Every Knowledge subdomain will contain one or more "elementary information". An elementary information can be implemented, increased and improved, depending of the necessity.

The principles of XBASE Tool are:

1) every Knowledge's system, although complex, can be decomposed in simple subsystems;

2) the efficiency of the process improve with the increase of the number of the information.

An example of the principle one could be the fractal geometry: in this case a geometric shape, although complex, can be decomposed in elementary and simple shapes.

The principle one could be explained in mathematical terms: given a Knowledge's system Ω, one can always find a decomposition in subsystems D_k:

$$\Omega = D_1 \cup D_2 \cup \ldots \cup D_k \cup \ldots \cup D_n \quad (2)$$

Every subsystem D_k is characterized by an elementary information χ:

$$\chi = f_c(\delta, \sigma) \quad (3)$$

with f_c is an elementary function, δ represents the list of input, ρ represents the list of output.

The function f_c is activated when the input δ takes on an admissible value. This function can be a fuzzy function, an heuristic function, etc. However, indicating with $\otimes$ the generic correlation between an input and an output, f_c is

$$f_c = \delta \otimes \sigma \quad (4)$$

So, the elementary information χ depends by the correlation between input and output.

Given a subsystem D_k, a list of input $\Im(\delta)$ the concept can be extended, examining all correlation between the list of input and a list of output $\Re(\sigma)$.

So:

$$[\Phi]m,n = \Im(\delta) \otimes \Re(\sigma) \quad (5)$$

with $[\Phi]m,n$ system of matrix.

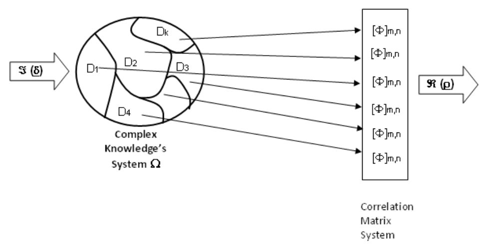

Figure 3: Correlation Matrix System

Whereas in a Knowledge system Ω, union of different subsystem D_k, every correlation matrix $[\Phi]m,n$ forms a system correlation matrix (figure 3). So, there isn't a single activation function, but a set of activation functions.

Starting with the list of input $\Im(\delta)$, through the correlation matrix system, it is obtained the list of output $\Re(\sigma)$, and so the results of problems.

So, every subsystem D_k produces a result σ, and if one considers the more sure results $\Re c(\sigma c)$, one can repeat the process, with a new level. In particular, at this level, the new input will be:

$$\Im'(\delta) \equiv \Re c(\sigma c) \quad (6)$$

In this case, one will have a new level of results:

$$[\Phi]'m,n = \Im'(\delta) \otimes \Re'(\sigma) \quad (7)$$

In this way, the number of input, and also of output, at second level is less than at first level. Step by step, if new

levels are created (2°, 3°, 4°…), the number of output will reduce and so, one can arrive to a correct and desired synthesis level.
XBASE Tool allows to develop cognitive networks that can be used in both numeric and textual Knowledge domain. It allows to develop Knowledge-Based Systems, in Real-Time for industrial processes, like environment monitoring (Mappa, Salvi and Tagliaferri, 1995).

3.2 Feature of XBASE and Software Platform

The framework of the Knowledge Base in XBASE is like a matrix, where every cell is an "Elementary Info" allocation. Each elementary info contributes to one or more output (metadata), according to an incremental knowledge composition.
So, XBASE allows a visual knowledge formalization and implementation by mean of the "elementary info" (a short functional I/O relation), to build a cognitive framework and an intelligent process control:
Data x [Knowledge] ⇒ Info (1°,2°,..)
Info x [expertise] ⇒ Decision Making
Since XBASE expert system contain heuristic knowledge, it must have a way of dealing with the problem of expressing and reasoning with uncertain data and inexact knowledge.
So, in the Knowledge Base of XBASE each Elementary Info combines itself with all the others correlated at different levels of abstraction, in a not deterministic way, so that a lack of this one info doesn't preclude the expert system reasoning functionality.
XBASE Tool has a software platform very perceptual, efficient and flexible. This tool allows to manager the company's Knowledge, autonomously and independently of informatics skills. XBASE Tool works like an expert, but it doesn't need of a human expert to work. In fact, to use this tool, it is necessary to imagine a system and design this system. It uses an holistic approach and both deterministic and not deterministic logics. It allows to obtain a Knowledge's synthesis, step by step, till to obtain the desired target information. XBASE Tool allows to explicit the manager's Knowledge, and to formalize this Knowledge. In this way, the user obtains a decision in automatic way, and he is conscious of the decisional process obtained.

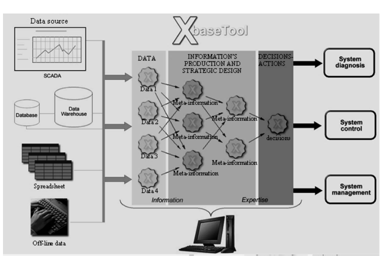

Figure 4: XBASE Tool's interface

The software's interface is very simple and perceptual: the user designs and moves the objects, similarly as he connects concepts in his mind.
In figure 4 is visualized an example of XBASE Tool's interface. The input data can come from a database, a spreadsheet and can be off-line data. The first level is made of these input data. At the second level, there are the meta-information, obtained by the correlation matrix system. So, it is possible a third level of meta-information (optional). Finally, there are the decisions-actions performing to solve the problem. It is possible to obtain the results in a very efficient way and in real-time.
So, XBASE allows to verify an initial manager's hypothesis and to obtain a simulation type "what-if". In this way, the manager can evaluate immediately the effectiveness of his hypothesis.

4. XBASE Tool: A Case Study

An interesting case study of XBASE's application is the on-Line/Real-time environment monitoring, for example monitoring of wastewater (Mappa, Tagliaferri and Tortora, 1997 **-** Mappa, 2003 – Mappa, 2000)**.**

4.1 The problem of the monitoring of wastewater

In this work an Expert System to on-line control water quality is presented, coming from textile finishing discharges, called BATTLE "Life" Project, Best Available Technique for Water Reuse in Textile SMEs.
The implementation of a BAT (Best Available Technique) for water reuse in textile finishing industries (in connection with environmental protection), requires to focalize a specific "Knowledge Base" around the on-line/real-time recognition of different "start of pipe" wastewater effluent patterns, to be segregated according to their suitability for treatment with membranes technology, belonging to the proposed BAT.
Textile wastewater is a mixture of many different chemical compounds and not all of these are recycled.
On-Line/Real-time monitoring of wastewater quality remains a scarcely resolved problem into the wastewater treatment industry.
A water quality assessment requires of more then fifty of specific chemical and physical parameters to be properly detected, all of them on samples by off-line laboratory instrumentation. Moreover, for water reuse scope (by using membranes technologies), on-line automatic controller and regulating flow valves require process real-time data-input.
A previous systematic characterization about quality and volume of the various process streams has been necessary, in order to identify those processes for which the substances contained in the various waste streams are still valuable and/or do not interfere with the quality of the product.
In spite of a deep changeability of textile segregated wastewater stream quality, there are only a few chances (instrumentation technologies) to on-line detect water

quality parameters and, no one of them, it is directly correlated to a specific pollutant.

4.2 "BATTLE" Basic Functional Model

To be able to solve the problem regarding on-line/real-time monitoring of textile wastewater quality, having only a few sustainable chances versus constraints satisfaction conditions, one suggests to resort to a multi-clustering on-line measurement technology (not multi-parameters only), based on WPR (Water Pollution Rate Index) indexes inferred to a appropriate Knowledge Base (network designed). In fact, to better detect and learn, from the on-line measurements and process data-input, to decide if an effluent has to be sent to the reuse plant or to the WWTP (how much fresh water has to be mixed with the permeate and all operative actions for the maximisation of reuse, in respect of the limits at the final discharge), an appropriate water pollution rate index has been defined and carried out, called WPR. This index has a mathematical normalized expression considering the main pollution driver patterns, apart from pH, previously balanced to a 6,5÷9 range value:

WPR = *f*(*dissolved salts; dissolved organic substance; total suspended solids*)$_{pH[6,5÷9]}$.

WPR will be able to on-line characterize the quality of wastewater streams and it will be "instructed" during the start-up phase (on the base of final effluent treatability evaluation and impact control, reclaimed water reusability, cost analysis, etc.) in order to be able to better control the selected streams. WPR is a value between 0 and 1, where WPR is close to 1, then the water is not good, but if WPR is close to 0, then the water is good.

4.3 The Expert System in "BATTLE" application

To face up to the produced water outflow variability, from printing and washing departments in particular, as well a not-sustainable direct and on-line measurement of referred chemical-physic parameters (DOC, TSS, VSS, turbidity, etc.), it will be used a real-time process controlled by Expert Systems technology (XBASE Tool), by realizing a prototypal version in this project context.

XBASE Tool will be used as a shift about the traditional process control systems (like SCADA), since they are not able, at present, to detect and to infer the not-direct measurements that actually it is sustainable monitoring (pH, temperature, conductibility, colour). This is instead of not measurable direct target parameters, as well in cognitive processes of diagnosis and decision-supporting elaboration.

The real-time Expert System water quality and recovery control will be operated by the following functionalities (figure 5):

• selection from Printing Department which of P1, P2, P4, P5, P6 water recovery stream is acceptable to be recovered on the base of WPR on-line measurement (pH, conductibility);

• selection from Washing Department which of P7-P8 water recovery stream is acceptable to be recovered on the base of WPR on-line measurement (pH, conductibility, colour);

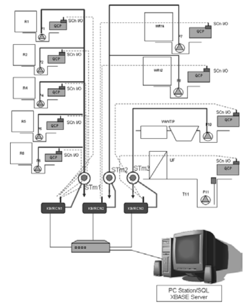

Figure 5: Functional Expert System scheme and its basic devices and components

• monitoring of UF plant by if P11 water recovery stream is acceptable to be treated by UF on the base of WPR on-line measurement (pH, conductibility, colour);

• monitoring of effluent from WWTP plant by if P10 water recovery stream is acceptable to be treated by UF on the base of WPR on-line measurement (pH, conductibility, colour);

• monitoring of Water Balance in the tank T11 so that, if necessary, ES can operates with the start of the related WWTP P10 pump;

• monitoring of Water Balance in the tank T11 so that, in the case of water recovered overflow, ES can operates with a more higher water quality limits.

4.3.1 Learning in "BATTLE"

In figure 6 is visualized the "core" of "BATTLE".

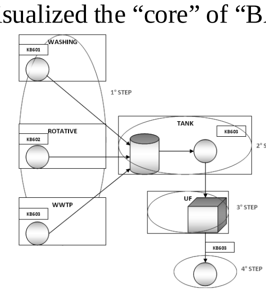

Figure 6: Principal steps of "BATTLE"

The principal steps are:

1) the water is collect in washing, printing and WWTP department, so it is analyzed by KB601, KB602 and KB603 nodes. WPR is extracted;

2) if water is good, it can be collect in tank;

3) the mixed water is analyzed by KB603 node: if this water is good, Ultra-Filtration process starts;

4) the water resulting by Ultra-Filtration is examined by KB603 node, to obtain an indicator about the final goodness of the water.

KB601, KB602 and KB603 nodes contains the Knowledge base to calculate WPR index, to value the quality of water.

This Knowledge base is codified, so that it can be altered simply by defining a vector of indices.
In particular, the Knowledge base is formalized through functions that represent the relations existing between the inputs received from sensors (pH, temperature, conductivity, color) and quality (WPR).
These parameters produce the vector K. The learning algorithms are introduced in order to obtain the vectors of indices K that realize the Knowledge bases, contained in the nodes.
In particular, there is an operator: this operator expresses a valuation of the water's quality, after points 1) – 3) and 4).

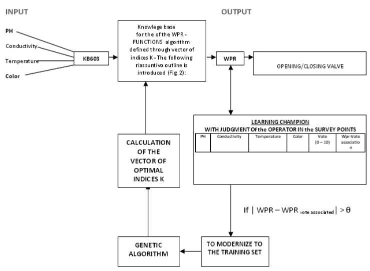

Figure 7: Members for the WPR calculation

The algorithm calculates vectors of optimal indices, so that the difference between the theoretical WPR and the detected WPR will be smallest.
The strategy of learning, chosen in BATTLE, takes advantage from Genetic Algorithms. Genetic Algorithms (GA) are procedures finalized to the optimization of the problem solving. They are based conceptually on the principles that regulate the natural evolution of the species. The GA simulate the evolution of a population of individuals: they support the evolution of optimal solution of a problem, so as in the human evolution process. The main steps of the physical process are:

• Generation of an initial population of individuals;

• Generation of new individuals, through a reproduction mechanism;

• The individuals compete each other and the winners have a greater probabilities to adapt to the environment;

• The population evolves, of generation in generation.

So, the problem is the search of the point of maximum of a function, called fitness function. For a specific problem, optimal solutions are selected, and later these solutions are combined to produce a final optimal solution. The GA promote the survival of the best solutions.
The Genetic Algorithm in BATTLE, as learning strategy, allows to obtain the optimal vector:

K = (K – PH, K – Cond, K – Temp, K – Color)

This vector influences the parameters PH, Conductivity, Temperature, Colour, and so WPR.
The Genetic Algorithm starts with the initial value of K - PH, K - Conductivity, K - Temp, K - Colour. Later, it tries to optimize the K-vector, in order that K calculates the WPR nearer to the votes of the operator: these new values minimize the difference between theoretical and detected WPR. Finally, the new K values will be visualized.

5. Conclusion

For the correct management of the great mine of data that there is at disposal, it needs a tool that allows to dispose in an optimal way the information, extracted by these data. This tool should help to take strategic decision, to solve every kind of problem. This should be realized in a prompt and automatic way, possibly in real-time, in order to optimize the working time.
So, rather then a complex and powerful computation systems, to a better "Problem Solving" approach, one makes use only of a good Knowledge Engineering ability and the right conceptual tools.
In particular, it has been realized a tool, XBASE Cognitive Mapping Tool. XBASE Tool allows to develop intelligent grids that can be used in both numeric and textual Knowledge domain. In this intelligent grid every node is an elementary information, and it forms the Knowledge Base of the examined problem. Every Knowledge Base is homogeneous, and the interconnections among the nodes allow to fuse different cognitive processes. This tool can be used in different Knowledge's domain, also in environment monitoring, as in monitoring of wastewater. In this area, it has been realized "BATTLE", a real-time process controlled by Expert Systems technology, XBASE Tool.

References

W. K. Cheung, J. Liu - "On Knowledge Grid and Grid Intelligence: a survey" - Computational Intelligence, Volume 21, Number 2, 2005.
J. Hia, N. Zhong, - "Organizing multiple data sources for developing Intelligent e-business portals" - Computer Science, Volume 12, Number 2, 2006.
G. Mappa, R. Tagliaferri, D. Tortora - "On- line Monitoring based on Neural Fuzzy Techniques applied to existing hardware in Wastewater Treatment Plants" - AMSEISIS' 97 - INTERNATIONAL SYMPOSIUM ON INTELLIGENT SYSTEMS - September 12, 1997.
G. Mappa - "Expert Software tools for Unfailing Water Quality" - TNO Environmental, Energy and Process Innovation - Apeldoorn, 21 March, 2003.
G. Mappa, G. Falivene, M. Meneganti, R. Tagliaferri - "Fuzzy Neural Networks for Function Approximation" - Proceedings of the 6th International Fuzzy Systems Association World Congress IFSA 1997.
G. Mappa - "Distributed Intelligent Information System for Wastewater Management Efficiency Control" - Wastewater Treatment Standards and Technologies to meet the Challenges of 21st Century 4-7th April 2000 AD - Queen's Hotel, Leeds, UK.
G. Mappa, et Al. - "On- line diagnostic system with intelligent software instrumentation based on neural fuzzy network" SMI 97 - Salone della Manutenzione - Fiera di Bologna, February 25, 1997.
G. Mappa, G. Salvi, R. Tagliaferri, - "A Fuzzy Neural Network for the On-Line Detection of B.O.D." - Wirn Vietri '95, VII Italian Workshop on Neural Nets ITALY, 1995.