Proceedings of ASCE 2005 International Conference on Computing in Civil Engineering, Cancun, Mexico; July, 2005

HUMAN-CENTRIC INTELLIGENT SYSTEMS FOR DESIGN EXPLORATION AND KNOWLEDGE DISCOVERY

I.C. Parmee[1]

ABSTRACT

This invited paper discusses research and development relating to computational intelligence (CI) technologies comprising powerful machine-based search and exploration systems that can generate, extract, process and visualise high-quality information from complex, poorly understood design domains. The integration and capture of user experiential knowledge within such systems in order to stimulate, support and increase understanding is of particular interest. The manner in which appropriate user interaction can overcome problems relating to poor design representation within systems utilising evolutionary computation (EC), machine-learning and software agent technologies is investigated. The objective is the realisation of user-centric intelligent systems that overcome initial lack of understanding and associated uncertainty; support an improving knowledge-base and stimulate innovation and creativity.

KEYWORDS

Intelligent systems, data-mining, visualisation, knowledge discovery and innovation

[1] ACDDM Lab, CEMS, University of the West of England, Bristol, BS16 1QY; Phone: ++ (0) 117 328 3137; Fax: ++(0)117 328 2587; ian.parmee@uwe.ac.uk

INTRODUCTION

Uncertainty and poor problem definition are inherent features during the early stages of design. Immediate requirements for relevant information to improve understanding can be confounded by complex design representations comprising many interacting variable parameters. Design constraints and multiple objectives that defy complete quantitative representation and therefore require a degree of subjective user evaluation further inhibit meaningful progression. In particularly complex areas, problem representation may, in the first instance, be merely based upon qualitative mental models arising from experiential knowledge, design team discussion and sparse available data. However, such representations, coupled with user intuition, play a significant role in defining initial direction and further investigation i.e. initial concepts based upon current understanding require both quantitative and qualitative exploration to generate relevant information that supports and enables meaningful progress.

PROBLEM REDEFINITION AND REFORMULATION

Generally, the development of computational design representations support exploration through the evaluation of solutions against criteria perceived to be relevant. Initial representations based upon current understanding and any available relevant data will likely be relatively basic and user-confidence in the model output may therefore be low. However, such representations can provide essential design insight despite their apparent shortfalls. Seemingly high performance solutions identified in terms of quantitative criteria followed by qualitative human evaluation utilizing experiential knowledge and intuition provides an indication of concept viability and model fidelity. An iterative user / machine-based exploratory process can commence where gradual improvements in understanding contributes to better representations, a developing knowledge-base and the eventual establishment of computational models that support more rigorous analysis. A highly interactive process thus emerges supporting the development of representation through knowledge discovery. Such a human / machine-based development may run concurrently with, and be enhanced by, other forms of investigation and of data / information gathering.

A high degree of assumption, particularly relating to objective representation, generally provides a starting point for our investigations. An initial variable parameter set may be selected with later addition or removal of variables as the sensitivity of the problem to various aspects becomes apparent. Constraints may be treated in the same way with the added option of softening them to allow exploration of non-feasible regions. Included objectives may change as significant payback becomes apparent through a re-ordering of objective preferences. Some non-conflicting objectives may merge whilst difficulties relating to others may require serious re-thinking with regard to problem formulation. The initial design space is therefore a moving feast rich in information (Parmee 2002). Such information, when extracted and coupled with the investigators' experiential knowledge and intuition, supports significant problem insight and subsequent problem re-formulation. It is quite possible therefore, that final solutions will be identified from a space that bears little resemblance to the problem space that provided a starting point for our investigations.

An example of design space reformulation could perhaps relate to the initial whole-system design of a major civil engineering project such as a large scale hydropower scheme where initial considerations relating to the major components (e.g. best local site, dam and spillway type, head and tailrace configurations, powerhouse position, access routes etc) will initially be dependant to a varying extent upon data relating, for example, to rainfall and run-off data, varying ground conditions, requirements of the client concerning mode of power supply, material and plant requirements etc. Although availability of information may increase on a daily basis initial design decisions will likely depend upon best possible assumptions based on current data, relatively basic whole-system design simulations and associated analysis. This would certainly be the case during pre-feasibility and feasibility studies. Rapid design reformulation becomes a major requirement as more definitive information becomes increasingly available from various human and machine-based sources and design team knowledge relating to the whole system becomes more established.

Alternatively, design uncertainty may relate to lack of knowledge of physical phenomena. Complex fluid flow situations, ground conditions that involve materials that defy definitive analysis, the structural analysis of complex architectural forms or the inclusion of state-of-the-art materials for example may all demand investigative work involving initial computational models that at least partially define the perceived problem. Increasing understanding, knowledge discovery and the resulting development of more definitive models eventually leads to the identification of design solutions that appear to best satisfy requirement.

Engineers in general are facing increasingly complex challenges relating to, for instance, more demanding environments, scientific development and constant competitive improvement in product performance. Pushing the envelope is becoming standard practice and computational systems that support knowledge discovery and innovation during the early stages of design are now essential.

During early design stages we could be considered to be concurrently negotiating two design spaces i.e.

1) The machine-based quantitative space that is bounded and inflexible when considered stand-alone (i.e. the space defined by all possible variable combinations of a computational design model). Search and exploration algorithms utilizing machine-based criteria representations to evaluate solutions can rapidly provide novel information from this space that aids problem understanding at a human level. Such understanding and subsequent search space redefinition can radically alter the initial bounds.
2) The investigators' mental representations of the design problem. These representations are only bounded by current knowledge and understanding. The development of this problem space relies upon external stimuli that includes the output from machine-based representation plus human intuition and judgement at both a quantitative and qualitative level.

The indication from previous conceptual design work is that the appropriate melding of these two spaces will support a holistic, knowledge-based approach that can result in significant step changes to machine-based objective representation and in overall understanding. We could consider all of the above to be a general description of how we progress when faced with poorly defined design problems that initially seem beyond our perceived analytic capabilities. Using this description the following sections explore a human-centric utilisation of evolutionary computation, machine learning and agent-based approaches integrated with enabling computational technologies to significantly enhance this knowledge discovery and representation development process. Particular areas requiring attention are:

- the development of meaningful computational design representations from experiential knowledge, sparse data and collective reasoning;
- non-linear search and exploration processes that can negotiate the complex solution spaces described by such representations;
- the capture of user experiential knowledge and intuition during re-definition of machine-based representations and reformulation and subsequent exploration of innovative solution spaces;
- development of software agent-based activities for information extraction, processing and succinct presentation to the user resulting in a reduction of cognitive load.
- The development and integration of machine-learning processes that support semi-autonomous activity.

The overall objective is the establishment of user-interactive computationally intelligent search and exploration environments that support rapid concept formulation, exploration and evaluation. Novel human-centred computational design processes should lead to innovation and competitive product development through continuous knowledge discovery.

Unfortunately, it is only during the later stages of the design process that sufficient data / information is available to satisfy current computer-aided design tool requirements. The earlier stages of design where innovation can result in very significant improvements remain very poorly supported by much of the powerful computational capability available. It is time to redress this imbalance.

RELATED RESEARCH

The concept of problem formulation and reformulation is well established within the design research community especially when considering innovative and creative design (Gero 1994, Goel 1997, Su 1990). This is associated with the development of a designer's understanding of a problem during the early investigative stages that may result in radical changes in problem representation. Another concept relates to the integration of knowledge from other sources through, say, analogical or metaphorical transfer from another problem area which can be of significant benefit especially with regard to the development of innovative approaches (Gero and Shi 1999; Brown 1998).

COMPUTATIONAL INTELLIGENCE TECHNOLOGIES

Computational intelligence (CI) techniques relevant to and developed within the design domain are now reaching a level of sophistication that allows them to be utilised to support a more holistic approach to problem solving. Machine-based exploratory systems can better handle the complexities of high-dimensional space ensuring that succinct specific information is available to the investigator thus enabling a greater user-concentration upon the significance of emerging results. These technologies can contribute in a piecewise manner to the development of user-interactive intelligent systems that meld machine-based design processing and user problem solving.

Much of the author's research has concentrated primarily in this area of the utilisation of CI techniques for the generation, extraction and visualisation of high-quality design information from complex design spaces. Early research addressed the identification of regions of robust high-performance solutions through the shared near-neighbour clustering (Jarvis and Patrick 1973) of genetic algorithm output and the subsequent analysis of solutions within each identified cluster (Parmee, Johnson and Burt, 1994). This work contributed significantly to the development of the cluster-oriented genetic algorithm (COGA) which has since provided the means to extract wide-ranging information relating to appropriate variable ranges, solution and variable sensitivity and the degree of conflict between included objectives and constraints (figures 1 and 2) (Parmee 1996; Parmee and Bonham 1999, Bonham and Parmee 1999). Near-neighbours clustering of GA output again played a role in Roy and Parmee's work relating to the identification of diverse high-performance design solutions that are subsequently passed through a fuzzy inference process. Here, a fuzzy rulebase generates a qualitative fitness assessment of the quantitative solution input providing a more relevant indication of a design's overall suitability (figure 3).

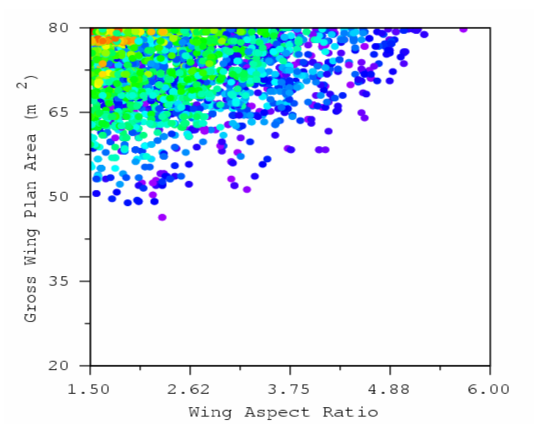

Figure 1: COGA single objective output showing the projection of high-performance (HP) solutions onto a two-variable hyperplane of a complex problem relating to preliminary air-frame design

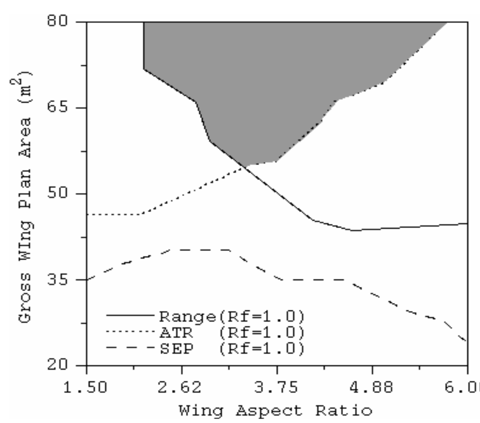

Figure 2: COGA air frame design output relating to 3 objectives again projected onto 2D variable space. This gives clear indications of degree of conflict between objectives (see Parmee and Bonham, 1999)

More recent work has further investigated the visualization of high-quality design information within interactive evolutionary design (IED) processes (Parmee et al 20001, Packham and Parmee 2000, Parmee 2002,). These user-centric processes have increasingly included the integration of software agent-assisted analysis of GA output that provides further support to the designer in the identification of complex relationships between variables, constraints and multiple objectives (Cvetkovic and Parmee 2001). On-line user-centric criteria ranking capabilities have been achieved via the integration of fuzzy preference techniques (Fodor and Reubens 1994, Cvetkovic and Parmee, 2003).

Further development of the COGA approach within interactive systems and the integration of data-mining and statistical analysis techniques has resulted in diverse graphical perspectives of objective / variable relationships and direct solution mapping between variable and objective space (figure 4). This mapping supports a far better understanding of the spatial relationships between high performance solutions that lie upon and close to a Pareto frontier (figure 5). The recent development of the Parallel Co-ordinate Box Plot (figure 6) gives an overall perspective of much of the information relating to variable and objective sensitivities contained in the two dimensional projections of figures 1,2,4 and 5 (Abraham and Parmee, 2004; Parmee and Abraham 2004).

Much of this research complements the well-established field of interactive evolutionary computation (IEC) which involves varying degrees of user involvement in the assessment of EC-generated solutions. Recent engineering applications of IEC include Takagi et al 1999, Carnahan and Dorris 2003 and Caleb-Solly and Smith 2002. A positioning of the above IED work along an implicit / explicit interactive evolution spectrum can be found in Parmee 2004.

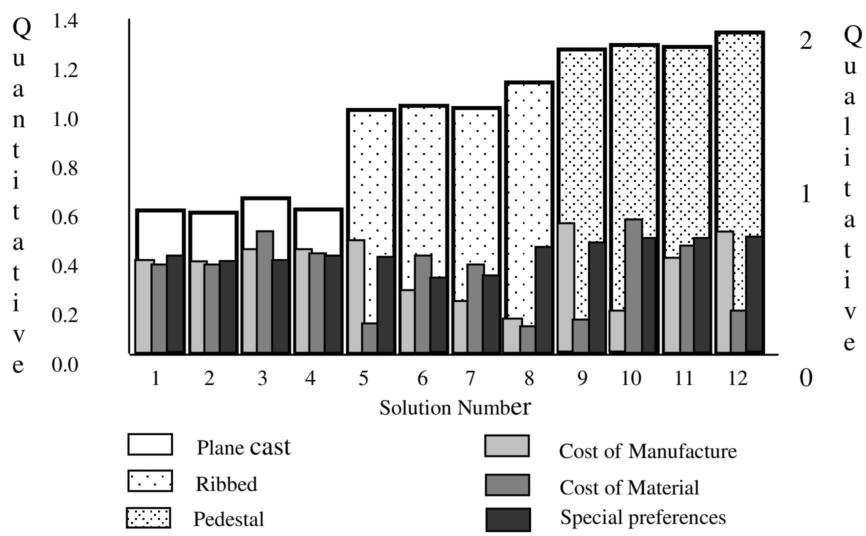

Figure 3: Bar chart showing twelve quantitative high-performance solutions relating to cooling channel performance that have been identified by a GA.. Quantitative solution fitness is depicted by the large outer bars. Qualitative measures generated from a fuzzy inference engine that applies qualitative rules to each GA solution are shown by the enclosed, smaller bars (see Parmee 2001).

Current related research within the ACDDM Lab is investigating the manner in which multi-agent systems can provide a machine-based negotiating capability that assists in solution identification and selection, qualitative objective satisfaction and the processing of search direction alternatives utilizing data extracted from evolutionary search and exploration systems (Abraham and Parmee 2005).

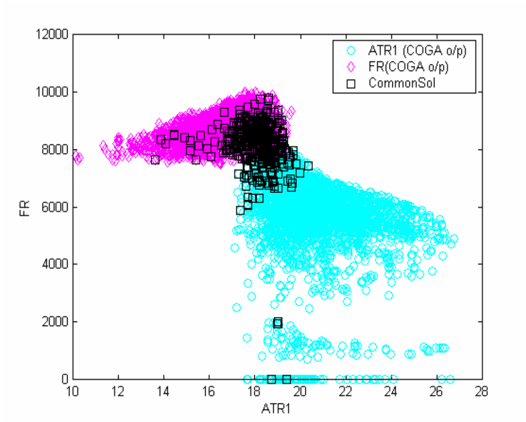

Figure 4: COGA projection of HP solutions relating to two objectives The blue cluster shows HP solutions relating the ATR1 objective whereas the pink cluster relates to the FR objective. The common black solutions satisfy both objectives (see Parmee and Abraham 2004)

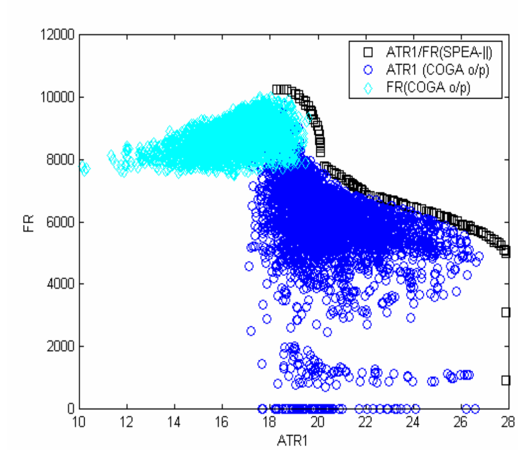

Figure 5: Similar projection illustrating how the edges of the cluster provide a close approximation to a Pareto frontier generated for the two objectives using the SPEA MOGA approach (see Abraham and Parmee 2004)

Other recent examples of design research involving evolutionary computation, visualisation and user-interaction include the concept generation work of Avigad et al (2004), Grierson's excellent work relating to the visualization of multi-dimensional Pareto frontiers (2002) and Packham's further development of clustering approaches relating to GA-generated solutions and the subsequent visualization of output (2004).

All of the above examples in addition to complementary computational intelligence (CI) developments in other domains supports the perception that these technologies could contribute in a piece-wise manner to a user-centric conceptual design environment.

Enabling Computational Technologies

The introduction of supporting and enabling technologies such as state-of-the-art data mining, statistical analysis, visualization techniques and high-performance computing (HPC) to such an environment would be essential. Appropriate integration would result in interactive CI search and exploration systems where the user becomes immersed within an information-rich computing environment accepting and analysing output and introducing change. High-performance computing capabilities would contribute to the achievement of a relatively seamless interface between interactive processes by reducing response time. On-line data-mining techniques (Hastie et al 2001) coupled with agent-assisted data processing and visualisation would greatly enhance the immersion concept. Overall integration with e-Science technology (Rana 2003) would lead to the establishment of Grid-based search and exploration capabilities that are widely available to the design community whilst also enabling remote access to very significant HPC resources and possibly to diverse information sources that enhance current knowledge of the problem at hand. Grid technology could also

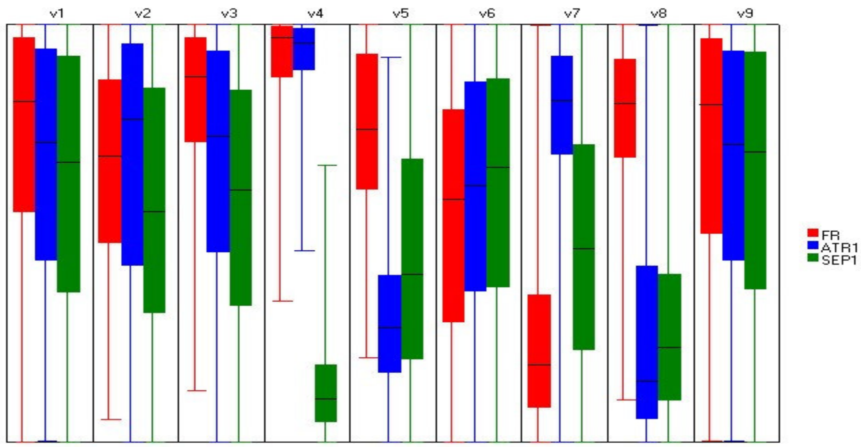

Figure 6: Parallel Co-ordinate Box Plot of high-performance solution distribution of each objective across all variable dimensions (V_n). The length of the three vertical axes related to each variable indicates to what extent COGA output for each objective covers each variable range. The degree of overlap of the three boxes indicates the manner in which each variable affects the degree of conflict between the objectives. A full description of this representation can be found in Abraham and Parmee 2004.

provide the means to avoid the development of monolithic systems. Computational components would be developed at remote centres of expertise and integrated with the design environment when required. Distributing problem solving systems supporting both statistical modeling, design space sampling and design search and exploration are currently under investigation within the ACDDM Lab (Parmee et al 2005).

THE ENVISAGED INTERACTIVE CONCEPTUAL DESIGN ENVIRONMENT

The establishment of a seamless user / machine-based design environment as described is ambitious. However highly efficient search across changing fitness landscapes with varying objective preferences and changing constraint conditions is achievable. It is also possible to spawn concurrent / complementary local search utilising appropriate algorithms. Constraint-handling techniques can be introduced that support exploration and allow information extraction relating to constraint sensitivity. Search space sampling techniques can be integrated with exploration processes to rapidly generate concepts of problem complexity as landscapes change. Statistical and CI-based modelling techniques are readily available and the concurrent utilisation of differing model types to provide better overall representation and increased confidence is accepted practice in some design areas. A possible configuration of the various system components and of user interactivity is simply illustrated in figure 7.

Taking this one step further, imagine developing relatively basic machine-based conceptual design representations associated with current understanding and then being able to rapidly explore the multi-variate space described by these representations using combinations of local and global search techniques. As search progresses the overall system

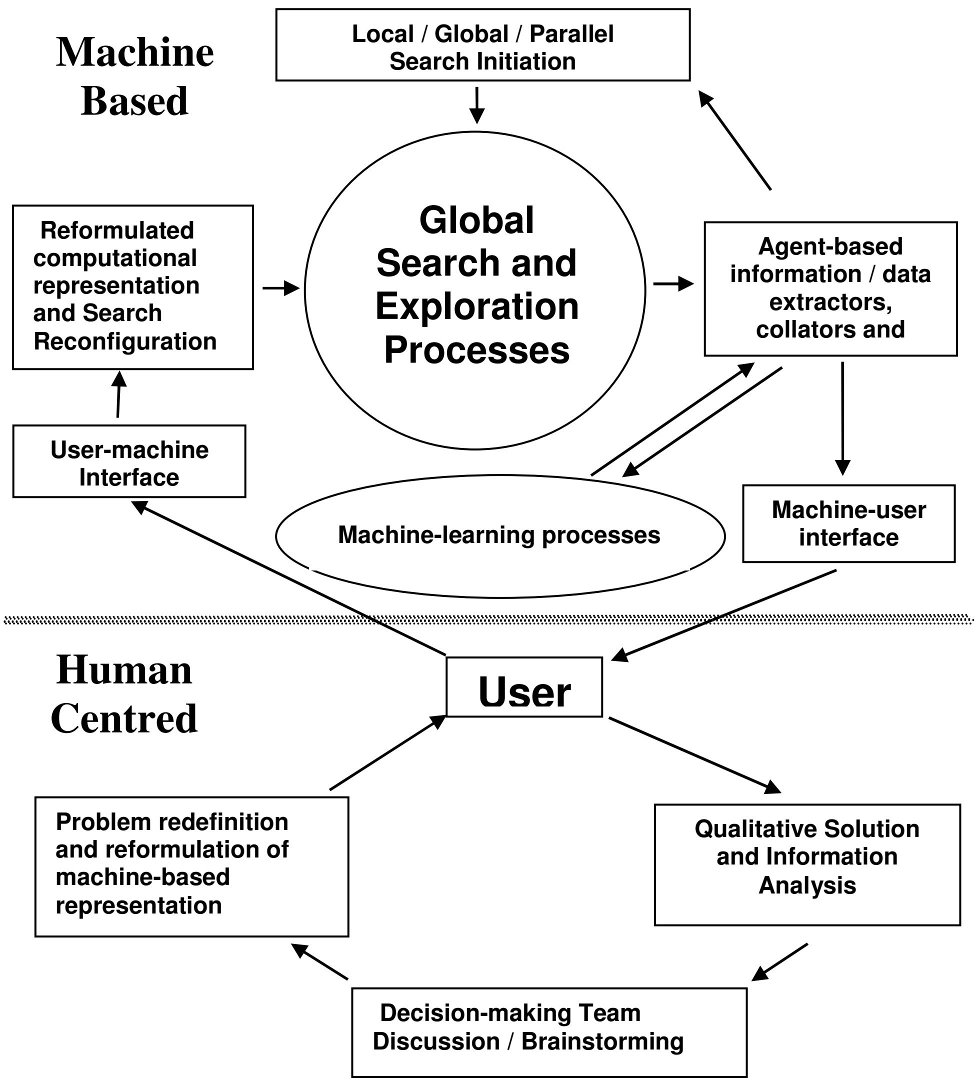

Figure 7: Simple illustration of a possible system configuration

is extracting and accumulating information relating to complex characteristics of the design domain whilst also discovering viable solutions. Solutions are initially identified that best satisfy objectives / constraints seemingly relevant in terms of current understanding whilst background processes extract information from areas of the problem space previously visited and present this, in a succinct manner, on-screen to the user. The degree of difficulty of satisfying initial objectives within existing variable bounds or within existing objective preference ranking becomes quantifiable and presentable through background data processing as search progresses. On-line user actions such as constraint softening, objective preference variation or modification of variable ranges may change the nature of the space and search direction whilst machine-based software agents acting as information collators, processors and presenters provide indications of the effects of such changes. These agents constantly advise the user on interesting solution correlation or re-direct you to previously visited areas now possibly of more interest. Concurrent, finer-grained, localised search processes may be spawned to explore specific regions. These actions become semi-autonomous as, through a machine-learning capability, the agents become more 'aware' of your requirements. The environment becomes more immersive as you react to the information being presented. User on-line actions become an integral part of the exploration process reacting to feedback from the system to make iterative changes to the problem landscape.

At any point this relatively continuous exploration process can be paused and relevant information downloaded and presented to the decision-making team for discussion. An easily understood graphic provides a recorded history of user-instigated change thereby supporting traceability and allowing analysis of the logical progression of the team's thinking based upon extracted information. The presentation of such material promotes discussion and allows the perspectives of others to be integrated in further exploratory interactive activity via appropriate problem re-definition and re-formulation.

As this iterative interactive process continues so confidence in the developing design models increases, the knowledge-base becomes well-founded and uncertainty significantly decreases. A natural result is a reduction in user-interaction as we move from a high-risk design definition phase through an intermediate phase of increasing confidence to the final stages of detailed analysis of a well-defined design space.

CONCLUSIONS

There is obviously much further research required to achieve the goal of the seamless user-centric system described above. However, many of the component parts are at a stage of development where their collective utilisation is possible and current research is pushing hard towards achieving this. It is suggested that the flexibility of CI technologies is such that specific problems are unlikely to be insurmountable. For instance, although a machine-based representation of an evaluation function may cause problems the user-centric approach supports complete or partial human evaluation of solutions and this can initially play an integral role.

Close interdisciplinary working will be essential to resolve arising problems. From an industrial point of view user-centric CI search and exploration systems could best utilise seemingly endless increases in desktop computational processing capability especially considering that in-house networked machines potentially support access to very high levels of distributed computing power. Such systems continuously running as background processes can support the development of in-house knowledge and expertise whilst reducing lead times to the discovery of innovative products when allied with complementary investigative processes. Current ACDDM collaborators in the pharmaceutical industry are already integrating our user-centric search, exploration and optimization processes using such in-house networked PC resource.

From an academic and industrial point of view the further development and utilisation of such systems within a research environment could support significant leaps in understanding relating to the characteristics of poorly defined complex design space. The ability to rapidly and efficiently play 'what-if' whilst concurrently gathering high-quality information that either confirms or contradicts current thinking suggests an environment well-suited to the support of knowledge discovery and innovation. The role of human intuition, experience and judgement within such an environment would be paramount whilst the inherent support of agent-based entities in terms of data processing and presentation would be invaluable.

REFERENCES

Abraham J. A., Parmee I. C. (2005) “Agent-based Systems for Mining, Processing and Visualisation of MOCOGA Output”. Procs International Conference on Engineering Design (ICED), Melbourne.

Avigad G., Moshaiov A., Brauner N. (2004) “Concept-based Interactive Brainstorming in Engineering Design”. *Journal of Advanced Computational Intelligence and Intelligent Informatics, 8(5).*

Bonham C. R., Parmee I. C., (1999a) “An Investigation of Exploration and Exploitation in Cluster-oriented Genetic Algorithms”. *Proceedings of the Genetic and Evolutionary Computation Conference*, Orlando, Florida, USA; pp 1491 – 1497.

Brown D. C., (1998) “Intelligent Computer-Aided Design” http://www.cs.wpi.edu/~dcb/Papers/EofCSandT.html; Revision of 1993 article for the Encyclopedia of Computer Science and Technology; J. Williams and K. Sochats (eds).

Caleb-Solly P, Smith J. (2002) “Adaptive Image Segmentation based on Interactive Feedback Learning”. Procs. International Conference on Adaptive Computing in Design and Manufacture, Spriger Verlag, pp 243 – 257.

Cvetkovic D., Parmee I. C. (2003) “Agent-based Support Within an Interactive Evolutionary Design System”. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing Journal*; Cambridge University Press, 16(5), pp 331-342

Cvetkovic D., Parmee I. C., 2001, “Preferences and their Application in Evolutionary Multi-objective Optimisation”. *IEEE Transactions on Evolutionary Computation*, **6**(1), pp 42-57.

Fodor J., Roubens M. (1994) “Fuzzy Preference Modelling and Multi-criteria Decision Support”. *System Theory, Knowledge Engineering and Problem Solving*, **14**; Kluwer Academic Publishers.

Gero J. S., Louis S. J., Kundu S., (1994) “Evolutionary Learning of Novel Grammars for Design Improvement”. *Artificial Intelligence for Design, Analysis and Manufacture*, 8, pp 83-94.

Gero J. S., Shi X-G., (1999) “Design Development based on an analogy with developmental biology”. CAADRIA ’99, J. Gu and Z. Wei (eds), Shanghai, China; pp 253-264.

Goel A.K. (1997) “Design, Analogy and Creativity”. *IEEE Expert, Intelligent Systems and their Applications*, **12**(3), 62 - 70

Goldberg D. E. (1989), *Genetic Algorithms in Search, Optimisation & Machine Learning*. Addison - Wesley Publishing Co., Reading, Massachusetts.

Grierson D. E., Khajehpour S. (2002) “Method for Conceptual Design Applied to Office Buildings”. Journal of Computing in Civil Engineering, 16 (2) pp. 83-103.

Hastie, T., Tibshirani, R. & Friedman, J. (2001) *The elements of statistical learning: data mining, inference and prediction*. Springer-Verlag, Berlin; 2001.

Jarvis R. A., Patrick E. A. (1973) “Clustering using a Similarity Measure Based on Shared Near-neighbours”. IEEE Transactions on Computers; 22[11].

Machwe A., Parmee I. C., Miles J. C. (2005) "Overcoming Representation Issues When Including Aesthetic Criteria in Evolutionary Design". Procs ASCE International Conference on Computing in Civil Engineering, Cancun, Mexico**.**

Packham, I., Parmee I. C. (2000) "Data Analysis and Visualisation of Cluster-oriented Genetic Algorithm Output". Proceedings IEEE Int. Conference on Information Visualisation, London.

Packham I.S., Rafiq M. Y., Borthwick M. F., Denham S. L. (2004) "Interactive Visualisation for Decision Support and the Evaluation of Robustness using Evolutionary Computing". 11th International EG-ICE Workshop, Weimar, Germany.

Parmee I. C., Abraham J. A., Shackelford M., Rana O. F., Shaikhali A. (1995) "Towards Autonomous Evolutionary Design Systems via Grid-based Technologies"

Parmee I. C. (2004) "User-centric Evolutionary Design". Procs. Design 2004, Dubrovnic.

Parmee I. C, Abraham J. R. (2004) "Supporting Implicit Learning via the Visualisation of COGA Multi-objective Data". IEEE Congress on Evolutionary Computation, Portland, pp 395 – 402.

Parmee I. C. (2002) "Improving Problem Definition through Interactive Evolutionary Computation". *Artificial Intelligence for Engineering Design, Analysis and Manufacturing Journal*; **16** (3), Cambridge University Press.

Parmee I. C. (2001) "Evolutionary and Adaptive Computing in Engineering Design". Springer Verlag, London.

Parmee I. C., Cvetkovic C., A. H., Bonham C. R., Packham I. (2001) "Introducing Prototype Interactive Evolutionary Systems for Ill-defined Design Environments". *Journal of Advances in Engineering Software*, 32 (6), Elsevier, pp 429 – 441.

Parmee I.C., Bonham C.R. (1999) "Towards the Support of Innovative Conceptual Design Through Interactive Designer/Evolutionary Computing Strategies". *Artificial Intelligence for Engineering Design, Analysis and Manufacturing Journal*; Cambridge University Press, 14, pp 3-16.

Parmee I.C. (1996). "The Maintenance of Search Diversity for Effective Design Space Decomposition using Cluster Oriented Genetic Algorithms (COGAs) and Multi-Agent Strategies (GAANT)". Procs 2nd International Conference on Adaptive Computing in Engineering Design and Control, Plymouth, UK; pp 128-138.

Parmee I. C., Johnson M., Burt S. (1994) "Techniques to Aid Global Search in Engineering Design". *Proceedings of International Conference on Industrial and Engineering Applications of AI and Expert Systems*, Austin, USA.

Rana O. F., Pouchard L. (2003) **"**Agent-based Semantic Grids: Research Issues and Challenges". *Journal of Parallel and Distributed Computing Practices.*

Takagi H., Ohsaki M., (1999) "IEC-based Hearing Aid Fitting". Procs International Conference on System, Man and Cybernetics (SMC '99) ,IEEE, Vol 3, 657 – 662,

Su N. P., 1990, *The Principles of Design*. Oxford University Press, New York.