doi:10.1093/comjnl/bxh160

Capsule Reviews

FAIROUZ KAMAREDDINE

The Capsule Reviews are intended to provide a short succinct review of each paper in the issue, in order to bring the content to a wider readership. This issue's Capsule Reviews were compiled by Fairouz Kamareddine. Professor Kamareddine is an Associate Editor of *The Computer Journal* and is based in the School of Mathematical and Computer Sciences at Heriot-Watt University, Edinburgh, UK.

Formal versus Material Ontologies for Information Systems Interoperation in the Semantic Web. ROBERT M. COLOMB
Various programs (or arguments) have been developed with the purpose of facilitating the interoperability among selected mutually aware applications/communities. For such interoperability to work well, a common ontology between the interoperating programs must exist. However, since there are several communities and many ontologies, and since one community may want to interoperate with a number of other communities, how do we deal with the different ontologies? This problem led to the upper ontology movement which advocates the development of an abstract description of what there is in the world in an application-independent form. This paper argues that attempts to support information systems interoperation by building catalogues of everything are doomed to failure. The paper shows that several upper ontologies would be expected to fail due to semantic heterogeneity. On the other hand, the paper argues that it is possible to follow Kant's proposal in his 'Critique of Reason' and to develop an application-independent subclass of upper ontology which can succeed in the face of semantic heterogeneity. Kant looked to the form of knowledge rather than to content. This paper follows Kant to find plausible constituents of a formal ontology which is independent of material content and which specifies not what there is in the world but how what there is can appear to information systems. In this way, a formal ontology can be used to construct a particular material ontology.

Atomic Hypermedia. DUNCAN MARTIN AND HELEN ASHMAN
Hypertext was first introduced by Bush in 1945 and then by Nelson in 1965 (the term hypertext was given by Nelson). Since then, hypermedia research has progressed in different directions. This paper lists some of this work and in particular focuses on the node terminology where an object encapsulates some amount of content. The problems of node-based hypermedia as explained in the paper include the referencing of composite nodes, the referencing within nodes and the limitedness of assigning a unique identity to each content area. The article's solution to these problems is to introduce 'Atomic Hypermedia' where the basic data structure is 'Atomic' and encapsulates multiple items of media within a single unified structure. An atomic data structure (ADS) consists of a number of atoms and may contain any number of dimensions (a dimension is an ordering of values held by a common attribute or property). An atom only exists in dimensions referred to by its properties. These properties can be adjusted freely. However, each atom must have a unique address (which is the set of its properties). The paper discusses how this ADS concept is sufficiently powerful to act as the basis of a hypermedia system where content can be referred to by reference rather than by copying. The paper presents how this can be achieved and then discusses the advantages of such an Atomic Hypermedia approach over traditional node-based structures.

Primitive Intervals verus Point-Based Intervals: Rivals or Allies? JIXIN MA AND PAT HAYES
In AI as well as in some other disciplines where temporal representations are crucial for the modelling of information, time has been represented by taking either time points or time intervals as primitives. It has been argued that in systems where points are taken as primitives and intervals are constructed from points, the Dividing Instant problem (an ancient puzzle where incompatible results occur at the boundary positions) might arise. Point-based solutions to this problem are not fully satisfactory. Although the interval-based approach bypasses the Dividing Instant Problem, it is inadequate for reasoning correctly about time change. Furthermore, each of the two approaches, (points as primitives or intervals as primitives) has its further advantages and disadvantages, and there is much debate on which approach is to be preferred. This paper provides a critical examination of both approaches and establishes that they can be represented as logically equivalent under a certain interpretation. In particular, the paper introduces

(i) the point & interval-based approach where both points and intervals are primitives;
(ii) the point & type-based approach where points and types are primitives and intervals are derived

and then studies the expressiveness of these two approaches (i) and (ii) discussing what happens to the Dividing Instant problem and the possibility of representing incomplete

temporal information. The paper shows that approach (ii) can overcome the disadvantage of the traditional point-based systems in representing incomplete temporal information and that it has all the expressive power of the interval-based approach. The paper then shows that both approaches (i) and (ii) are logically equivalent.

Caching and Perfecting Algorithms for Programs with Looping Reference Patterns. GIANLUCA DINI, GIUSEPPE LETTIERI AND LANFRANCO LOPRIORE

When a referenced page is not available in the primary memory, the memory management system fetches (using the secondary memory system) the missing page from disk to primary memory. The memory management system, depending on the situation, calls its page replacement algorithm. The least-recently-used (LRU) algorithm replaces the page whose previous reference is furthest in the past and has a relatively good performance. However, the LRU algorithm produces the worst possible performance in some cases which involve the iteration of a sequence of accesses (since it evicts pages before they are referenced again). In such cases, it is better to use the most-recently-used (MRU) page replacement strategy which evicts the page whose previous reference is nearest in the past. Furthermore, various fetch/prefetch techniques have been proposed to reduce storage access time keeping in mind that: (i) a prefetch started too early may cause the replacement of a page that will be used soon, and (ii) a prefetch started too late may cause the program to stall waiting for termination. This paper presents a thorough analysis of the memory behaviour of a number of caching and prefetching algorithms, with the main emphasis on the iteration of a sequence of accesses to a set of memory pages. Both the LRU and MRU page replacement algorithms are taken into consideration in this study, which together limit the range of the possible response times in the absence of page prefetch. Both the early prefetch and the late prefetch page prefetching algorithms are also considered. The analytical evaluation and comparisons of these four algorithms are given. A large set of measurement experiments is carried out.

Relational Formalism for the Management of Spatial Data. NIKOS A. LORENTZOS AND JOSE R. R. VIQUEIRA

Initial approaches to the management of spatial data have concentrated on the precise geometric representation of spatial data and on the implementation of operations on spatial data. This created many shortcomings including the lack of a good underlying formalism. Spatial databases attempted to solve such shortcomings; however, operations on spatial objects tended to be individually oriented rather than general, leading to various limitations. The present paper proposes a spatial data model and a relational formalism which attempt to overcome some of the limitations. In particular, the paper formalizes a spatial quanta which forms the basis for three spatial types (point, line and surface) which guarantee that every spatial object is a closed set. Predicates, functions and spatial operations such as difference, intersection etc., (on sets of spatial objects) are then defined. Again, the formalism guarantees that the result of a spatial operation consists of closed objects. Relational data structures and relational algebra operations (e.g. fold and unfold, quantum operations, pair-Wise operations and the like) are then defined. Many of these operations apply subset-wise and the authors realizing the importance of tuple-wise applications, study an extension which enables such tuple-wise applications. The functionality of the proposed model is demonstrated by means of various examples. Implementation issues as well as related work are discussed.

From Consensus to Atomic Broadcast: Time-Free Byzantine-Resistant Protocols without Signatures. MIGUEL CORREIA, NUNO FERREIRA NEVES AND PAULO VERISSIMO

Byzantine faults are arbitrary faults where arbitrary failure can occur. For example, processes can stop, omit messages, send incorrect messages, send several messages with the same identifier etc. Byzantine-resistant (or intrusion-tolerant) protocols which are capable of tolerating Byzantine faults are important for the construction of secure systems. However, Byzantine-resistant protocols have higher time and message complexity and use more CPU-time than crash-tolerant protocols. The paper presents a stack of three Byzantine-resistant protocols which avoid as far as possible the use of public-key cryptography (and are hence more economic on CPU-time), make no asynchronous assumptions, have a low-time complexity and an optimal resiliency. The paper starts from the fundamental problem of consensus in a distributed system which amounts to the question: 'how does a set of distributed processes achieve agreement on a value despite a number of process failures'. The three protocols are built on top of a randomized binary consensus protocol and a reliable broadcast protocol. The first protocol that the paper proposes is a multi-valued consensus where processes can propose values with arbitrary length. The second protocol proposed is a vector consensus which makes agreement on a vector with a subset of the values proposed instead of a single value. The correctness proofs of both the multi-valued consensus protocol and the vector consensus protocol are given. The paper then presents an atomic broadcast (or total order reliable broadcast) protocol which delivers the same message in the same order to all processes. The atomic broadcast protocol is implemented on top of the vector consensus protocol and it is shown to be correct. Equivalence results relating these three protocols are given and this is followed by an analysis of their performance.

Probabilistic Visual Cryptography Schemes. S. CIMATO, R. DE PRISCO AND A. DE SANTIS

A visual cryptography scheme for a set P on n participants is a method to encode a secret image into n shadow images called shares where each participant receives one share. Some

elements of the set P belong to a so-called 'qualified set' and are allowed to recover the secret image (without need for any knowledge of cryptography or having to carry out cryptographic computations) whereas others belong to a so-called 'forbidden set' and are not allowed to recover any information on the secret image. Usually, in the classical deterministic models, the secret image is reconstructed using a certain number of pixels. However, recently a new model was proposed by Yang where the reconstruction of new images is probabilistic. In this probabilistic method, there is no pixel expansion and hence the reconstructed image need not be n times (n being the number of subpixels used to reconstruct it) bigger than the original one. However, this comes at a price. Whereas in a deterministic scheme it is guaranteed that for any qualified set of participants, the pixel is reconstructed correctly, in a probabilistic scheme, there is a (small) probability of mistakes occurring in the reconstruction of the secret image. This paper deals with this tradeoff between pixel expansion and the probability of a good reconstruction and introduces a model which can be seen as a generalization of both the traditional deterministic model and the probabilistic model. The paper provides studies of when a deterministic scheme can be transformed into a probabilistic one, and when a probabilistic scheme can be transformed into a deterministic one with pixel expansion. This study is used to construct probabilistic schemes with pixel expansion starting from a deterministic scheme, transforming it into a probabilistic scheme and then transforming it further into a probabilistic scheme with pixel expansion. These newly constructed probability schemes with pixel expansion satisfy the security property. Furthermore, a formula for the probabilities of these new schemes is provided as a function of the probabilities of the starting scheme. This construction is demonstrated by applying it to some well known deterministic schemes such as the (n, n)-threshold and the $(2, n)$-threshold schemes. For the (n, n)-threshold scheme, it is shown that there is a linear relation between the pixel expansion and the probabilistic factor.

Can Unbreakable Mean Incomputabe? ALEXEI VERNITSKI

A problem is said to be non-computable if it has been established that there is no algorithm which can solve this problem. A computable problem (i.e. one for which an algorithm exists which solves this problem) is said to be intractable if all the algorithms which solve it are inefficient. All modern commercially used cryptosystems are based on intractable problems in the sense that, the problems of breaking such systems are intractable. Such cryptosystems are unbreakable as long as the attackers do not have fast computers or cannot invent efficient algorithms to break them. The question of whether unbreakable encryption can be understood as the non-computability of breaking such encryption has occasionally been posed in the literature. This paper concentrates on this question within the context of one particular formalization of the concept of cryptosystems (many such formalizations exist in the literature). The paper shows that for this particular formalization, the problem of breaking the system cannot be made non-computable. In particular, the paper shows that in such formalization, cryptosystems which interpret 'unbreakable' as 'non-computable' do not exist. The paper concentrates on exhaustive key search attacks against a cryptosystem. The paper shows that for each cryptosystem in this formalization, there is a key search algorithm which leads to an exhaustive valid key search and that there are cryptosystems whose sets of valid keys are not recursively enumerable.

A Probabilistic Model for Information and Sensor Validation. PABLO H. IBARGÜENGOYTIA, SUNIL VADERA AND L. ENRIQUE SUCAR

A decision which is based on a faulty data could lead to disaster and hence, information needs to be validated and such validation must happen in real time. This paper presents a new theory and an 'any time' algorithm for probabilistic information validation. Any time algorithms are algorithms which can be interpreted at any point during computation and which return an answer whose quality increases as it is allocated additional time. The paper considers information to refer to the set of variables that some application requires for its decision process. The values of these variables can be provided by sensors. The sensor is said to be 'faulty' if the output measurement incorrectly represents the input. In this way, information validation amounts to the validity of the reading reported by a collection of sensors. Since decision making about information is imprecise in general, fuzzy approaches as well as neutral networks have been commonly used in the field of information validation. This paper opts for the use of Bayesian networks which enable the authors to make use of the significant amount of research on learning Bayesian networks, to use probabilistic propagation in order to estimate the probability distribution of a particular variable, and to capitalize on the fact that a faulty reading leads to the observation of potential faults in an important set of variables known as the Markov blanket. The validation algorithm developed in the paper is assessed by applying it to the validation of temperature sensors of the gas turbine at the Gomez Palacio power plant in Mexico. The results were very good with 95% accuracy for severe faults and 80% accuracy for mild faults.