paper discusses some trends in spatial database technology and implementation and the way in which these interact with the trends in update technology. To a large extent this interaction promises to be a beneficial symbiosis, but some problems still to be solved are identified and probed.

Some key developments in spatial database technology are described, including object-orientation and versioning. The impact of these on the utility of large spatial databases is considered along with further developments in client/server architecture, distributed processing and Internet access. The critical role of database integrity emerges, and this is examined particularly in the context of revision and update. The parallel tasks of information refinement and enhancement are also examined.

1.2 Data Transfer Standards

The relevance of emerging trends in data transfer standards is briefly addressed, including the transfer of update information. Short and longer term approaches to multiproduct databases, supporting both a range of cartographic products and a range of digital data products, are discussed with particular reference to update processes. Finally the view is advanced that by the Millennium the convergence effect that is already apparent in update technology, both from field sources and from imagery, will have resulted in such a degree of standardisation that there will be very little in the way of discriminating factors between rival offerings. The discriminating factor will be the database, and the ease and efficiency with which any particular update system can integrate with it.

2. TRENDS IN SPATIAL DATABASE TECHNOLOGY

2.1 Value in Information

Investors in spatial databases, whether they be governmental agencies or commercial companies, are being driven to increase the value of their information. In addition to refining its accuracy and currency they seek to add to its utility and range of application. Ideally they seek to do this without having an ever increasing proliferation of independent databases, all demanding costly maintenance. In practice, this aim is not immediately achievable. A viable approach is to reduce the number of independent databases, and to automate, or at least orchestrate, the posting of updates across them. Software tools which identify, validate and codify changes together with flowline management tools and data models which provide means of administering changes provide the essential framework. A conceptual model is shown in Figure 1.

2.2 Object-Orientation

Object technology provides a very productive approach to increasing the value of spatial information. Modelling of the subject is in terms of objects that closely fit the real world, rather than just in terms of geometries and relational tables. Object-orientation (O-O) goes a step further, and allows the behaviours of objects to be modelled. O-O databases and O-O programming languages are well-established in mainstream Information Technology (Informatics) and are being effectively deployed to handle large volumes of spatial information in a very flexible manner. In particular O-O spatial databases provide very efficient support for topology and spatial generalisation, and have very positive scaleability characteristics. They can provide large area support from a common base for both a range of map scales and a diversity of data products (e.g. link-node transportation networks and polygonised boundary data). An overview of Object-Orientation in the context of Geographic Information is to be found in (Woodsford, 1995) and a more popular treatment from a general management perspective in (Taylor, 1990).

2.3 Versioning

Database Versioning is a second crucial advance. It provides support for very large continuous spatial databases in an economical and manageable fashion. Versions are maintained in a tree structure, with only change information recorded, rather than complete copies. Multiuser update is supported, with each update process having its own logical copy of the complete dataset, without the overheads of providing a physical copy to each process. The concepts are illustrated in Figure 2, which also shows how versioning can be used in conjunction with long transaction support. An Update process (or user) has exclusive write access to a defined segment, which is simply a logical set of objects within the database. The set may be defined by a spatial extent, a set of object classes or any other logical rule. Segments have to be logically exclusive, although any process can have read access to the whole data.

Updates can be validated for internal correctness as they are generated, and validated for consistency with the database as a whole prior to merging the updated version into the master data. The mechanism for validation is the use of *methods*, which are the general mechanism in an O-O system for invoking behaviours. General and specific validation methods can be built into the object database schema (i.e. not at the application programme level) and used to ensure integrity of the data across update processes.

A powerful concept is that of the Object Lifecycle, illustrated in Figure 3. This provides a framework for managing all phases of the life of an object, including updates over time, by defining methods to be invoked at each relevant stage in transactions involving objects, and ensuring they are invoked. The mechanism is analogous to that of *triggers* in relational databases.

SPATIAL DATABASE UPDATE - A KEY TO EFFECTIVE AUTOMATION

Peter A. Woodsford
Laser-Scan Ltd, Cambridge, UK.

Commission IV, Working Group 3 - Map and Database Revision.

KEY WORDS: Updating Spatial Databases, Object, Standards, Object-Orientation, Versioning.

ABSTRACT:

The role of the spatial information database is central to both the efficient automation of mapping and the effective provision of geographic information. Dominant factors in the role of the database include the means of access, the services provided and the currency and validity of the data. Currency and validity depend on the update processes available and how they are implemented and controlled. The creation and preservation of data integrity is crucial.

Database technology is in a period of rapid advance. Object-orientation and database versioning are two techniques of particular value in handling large spatial databases, alongside trends in client-server architecture, distributed processing and Internet access. Update processes include both information refinement and enhancement as well as the incorporation of new information. Metadata, reflecting the status and quality of information, is key to managing database update.

There are parallels between managing incoming update processes (the reception of new information) and issuing updates to data holders. Emerging standards for data transfer fall into two broad classes - data-centric and process-centric. Both of these paradigms are relevant to the update of spatial databases and the issue of update information to users. The sources of update information (imagery, field data) and the associated mechanisms are becoming standardised. The discriminating factor is increasingly the database, and the means by which update processes can integrate with it. The balance of advantage is shifting towards performing update in the database environment, rather than by stand-alone processes.

RÉSUMÉ

Le rôle d'information spatiale de la base de données est au centre de l'automatisation efficace pour faire les cartes et aussi la provision effective de l'information géographique. Les facteurs dominants dans le rôle de la base de données comprennent le moyen pour accéder, les services fournis et le cours et validité des données. Le cours et validité dépendent sur les procès disponibles pour les mettre au jour et comment ils sont exécutes et contrôlés. La création et conservation des données est crucial.

La technologie des bases des données est dans une période d'avancement rapide. L'orientation d'objet et le versionement des données sont deux techniques d'une valeur particulière pour ma nager des grands base de données spatiale, a cote des tendances dans l'architecture du client-serveur, des procès distribues et de l'accès a l'Internet. Des procès pour mettre au jour comprennent le raffinement et l'utilisation rationnelle de l'information aussi que incorporation de nouvelle information. Metadonnés, qui reflètent la situation et qualité de l'information, c'est la solution pour gérer le mise à jour des données.

Il y a des comparaisons entre gérer le mettre au jour des procès des données en entrée (réception de nouvelle information) et la distribution des mise à jour des gardes de données. Les standards qui émergent pour transférer les données tombent dans deux large classifications - données-centré et procès-centré. Toutes ces paradigmes sont relevant pour mettre au jour des bases de données spatiales et dl'information (images, données de camp) et les mécanismes associes qui deviennent standardises. Le facteur discriminant c'est de plus en plus la base de données, et le moyen par lequel les procès de mise à jour peuvent entre intégrés aux données. Le balance de l'avantage se déplace vers l'exécution de mise au jour dans l'environnement des bases de données, plutôt que par des procès indépendants.

1. INTRODUCTION

1.1 The Role of the Database

Many factors contribute to the degree of success achieved in the automation of mapping and the effective provision of geographic information. It is becoming increasingly evident that the dominant factor is the role of the spatial database, the means by which it can be accessed and the services it provides. A significant aspect of this role is the degree to which the information in the spatial database is validated and maintained. The

Validation methods can be used to confirm local data integrity (i.e. within the current segment). Merge methods are particularly important as they confirm global integrity before the final commit at the end of a long update transaction. Note that caching is used to rollback any transaction that fails validation checks prior to the final commit.

Since the version holds an explicit record of all changes made in the transaction, the merge method can readily summarise the changes for the purposes of management information and for transmission (in digital form) to other update processes. The merge methods can also update any relevant metadata records within the database, to record the status of any changes made, and create any history objects that may be required to enable recover of previous states of the data. In short, merge methods, as well as ensuring overall data integrity, play a crucial role in the overall information flows shown in Figure 1.

3. TRENDS IN UPDATE TECHNIQUES

3.1 Systems

With the advance of digital techniques, a general paradigm for the update process is becoming established, based on the display of existing information against a backdrop of new source information, or selections of new source information. This is achievable using standard workstations or personal computers (soon even the most demanding image interpretation and stereo information extraction tasks will be accomplished without recourse to special hardware). These hardware advances are paralleled by advances in software engineering such as those supporting co-operating processes, and the strong move to Open IT environments. Of recent note is the strong take-up of the GEOTIFF standard, to provide workable registration of imagery data.

3.2 Data Models

The update process in this paradigm involves display of and interaction with all the relevant current data i.e. not just geometry and attributes, but topology, metadata and higher level relationships. The more success there is in increasing the information content (value) of the data model, the more the update process has to interact with that data model to maintain its integrity. In practice, the evolution of the data will involve not just update (reflecting real world change), but also accuracy refinement (and hence metadata update). In many cases it will also take place against a background of "data re-engineering", as progressively more complexity is added to the data model to reflect the demands to add value. The life-span of the data already exceeds that of the hardware and software systems on which attention usually focuses, and this trend will accelerate.

The evolution of the data model must be managed in carefully controlled stages, hopefully few in number ("one giant step?"). A key element within this is the management of the evolution of the update process, and the provision of sufficient metadata records for tracking purposes. The appropriate amount of such metadata is still a subject of concern and debate. The balance has to be struck between fitness for purpose, and the possibility of overwhelming the data with metadata. The same issues arise in considering how much history data to retain. As pointed out at the end of section 2.3 above, the technology now exists to retain records of all changes in such a manner that previous states of objects can be restored. It may of course be uneconomic or unnecessary to do so!

3.3 Orthoimages

Against this background, the full significance of digital ortho-images can be seen - as digital source documents to be assessed and referenced along with other digital source documents by processes that interact with and update the database. Their value lies in the ease with which they can be handled by such processes - they demand no special status and fall in naturally with the evolution of the data model and the database integrity maintenance mechanisms. To similarly exploit the benefits of richer sources of imagery (e.g. stereo), the interpretation process has to be closely coupled with the database. This is in principle possible and is sometimes achieved, but requires an open architecture on both sides (or a closed proprietary system covering both aspects!). There may of course be perfectly valid operational reasons for retaining the update process detached from the database (e.g. it may be government policy to contract out some of the activity), in which case the issues arising have to be addressed. These in fact broadly mirror the issues that have to be solved in managing the release of updates to users.

4. TRANSFER STANDARDS AND OPEN ARCHITECTURES

4.1 Data Exchange and Incremental Update

A number of developments are occurring to respond to the issues raised above. In some user communities, it is possible to propound a single data model of sufficiently wide utility. Such data models are usually object-based. Transfer standards can be defined, both for data exchange and for incremental update. Incremental update hinges on a working scheme for unique object identification ("object-ids"). In such a situation, it is quite feasible to extract a portion of the database for detached update, although integrated update is likely to be more efficient. The key technical challenge is to devise workable schemes for unique object-ids, particularly where there are multiple issuers/owners of data. Solutions have been proposed, but are not yet visible at the working level.

However the restrictions arising from such a centralised approach to the data model are onerous for a wider user community. Recent work on data transfer standards (as exemplified by the draft European (CEN TC287) transfer standard) have tackled the issue of transferring the data schema as well as the data. If taken up and successfully implemented, this approach would represent a significant step forward. In the update context, it would potentially make detached update more viable, as transfer of the data model (provided the update process can accept and use it) should allow the update process to update the data in terms of the model, performing the necessary integrity checks before returning the updated data to the master database.

4.2 Data-centric v. Process-centric

Alongside the progress towards "data-centric" transfer standards cited above, there is growing move towards "process-centric" standards. The motivation is to provide mechanisms for access to data where it sits, in its "local" format. Powerful stimulus for this approach is coming from the Internet community and the wider IT industry. The debate in the formal standardisation bodies (ISO TC211, CEN TC287) is still in progress, and the outcome is presently unclear, although the tide is moving in the process-centric direction.

The same trend is apparent in the Update process, for the reasons stated in this paper. The need is for the data model to be accessed via the Application Programme Interface (API) of the spatial database. Object-orientation can provide a very clean interface at this level, particularly in terms of maintaining spatial database integrity, which is the increasingly dominant concern.

The parallels between the activity of receiving and managing update information and the activity of issuing data and updates are apparent in the trends in technology and standardisation.

5. CONCLUSIONS

These trends are rapidly leading to a point at which the ability of update processes such as soft copy photogrammetry to deliver coordinate information and thematic interpretations is increasingly standardised, and in effect determined by the nature (resolution etc) of the available imagery rather than the processing available in the update tools. The utility of the update tools will rather be determined by their abilities to interact with increasingly complex data models, either in a detached mode via transfer formats that carry data model as well as data or in direct mode via the Open APIs of spatial databases with powerful data modelling capabilities. Increasingly these are following the Object-Oriented paradigm.

Within this framework, progress can also be anticipated on spatial generalisation, in the short term by orchestrating the posting of updates between databases supporting different scale ranges. As the powerful conceptual framework of Object-Orientation provides more usable spatial generalisation functionality, increasingly scale-independent databases will become possible (e.g. across a band of scales), but update of these will involve interactions with the underlying data models. These models must include provisions for metadata, and some appropriate mechanisms for recovery of history. The issues are trade-offs of system cost against data utility.

As spatial databases cover wider areas and have richer content, so the requirement increases to issue data only to those entitled to it, to issue changes in a timely manner and to validate both content and updates. There is a growing need to devise means of managing change only (or incremental) update. Just as the balance has shifted from primary data capture to the implementation of updates to data already held, so it will shift further to the issuing of updates to the users or customers for the data. There are underlying common factors to both the receipt and the issuing of change information. A key concept is that of unique object-ids. Object-Oriented data models and database versioning provide powerful techniques for Spatial Database Update - a Key to Effective Automation.

References from Books:

Taylor, David A. (1990). Object-Oriented Technology: A Manager's Guide. Addison-Wesley Publishing Company.

References from Other Sources:

Ibbs, Tony J. (1995). Progress in the standardisation of geographic data - a U.K., European and International Perspective. IUSM Working Group on GIS/LIS, Hannover, September 25-28, 1995.

Woodsford, Peter A. (1995). The Significance of Object-Orientation for GIS. IUSM Working Group on GIS/LIS, Hannover, September 25-28, 1995.

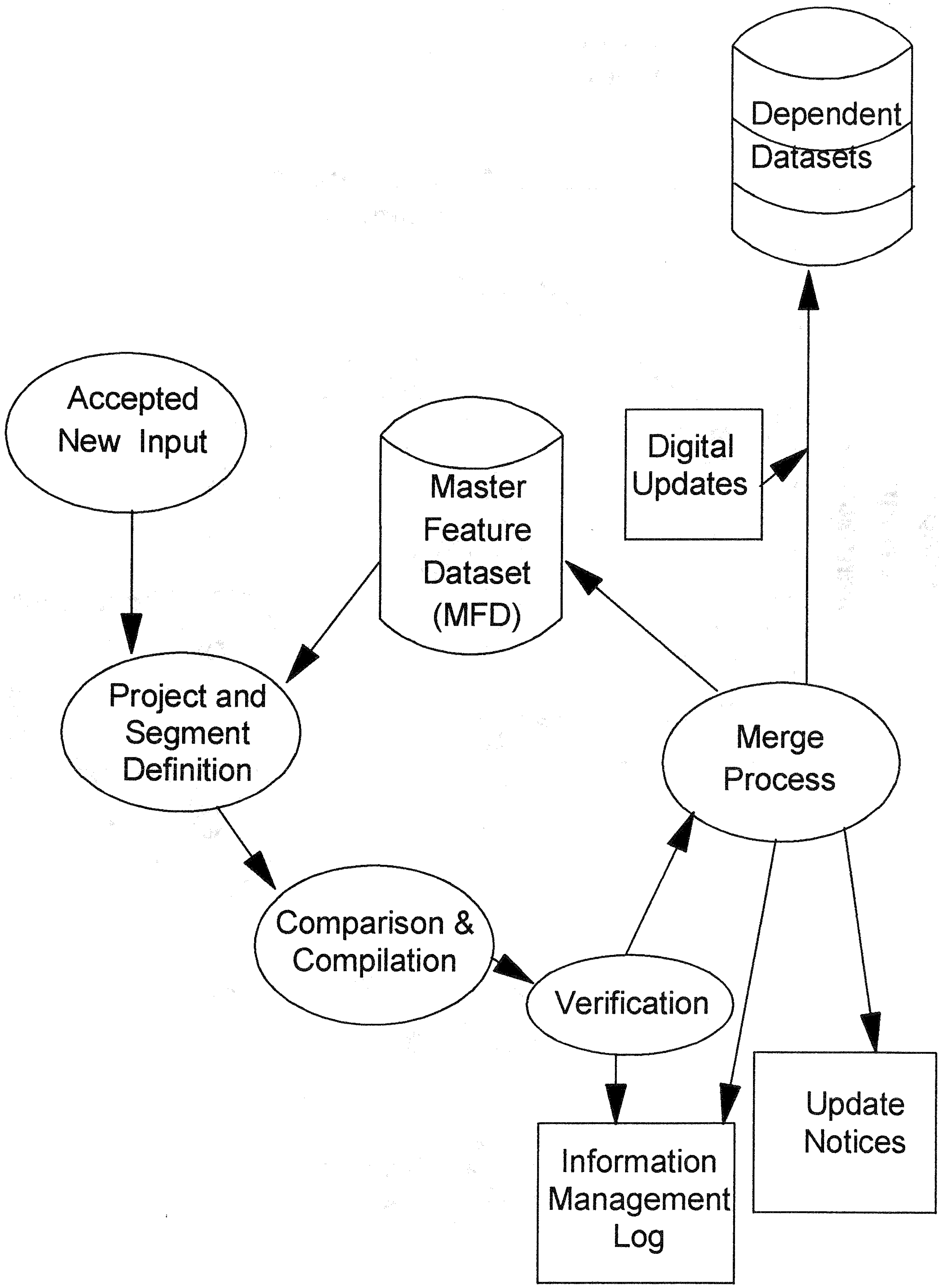

Figure 1: Information Flow and Update Processes

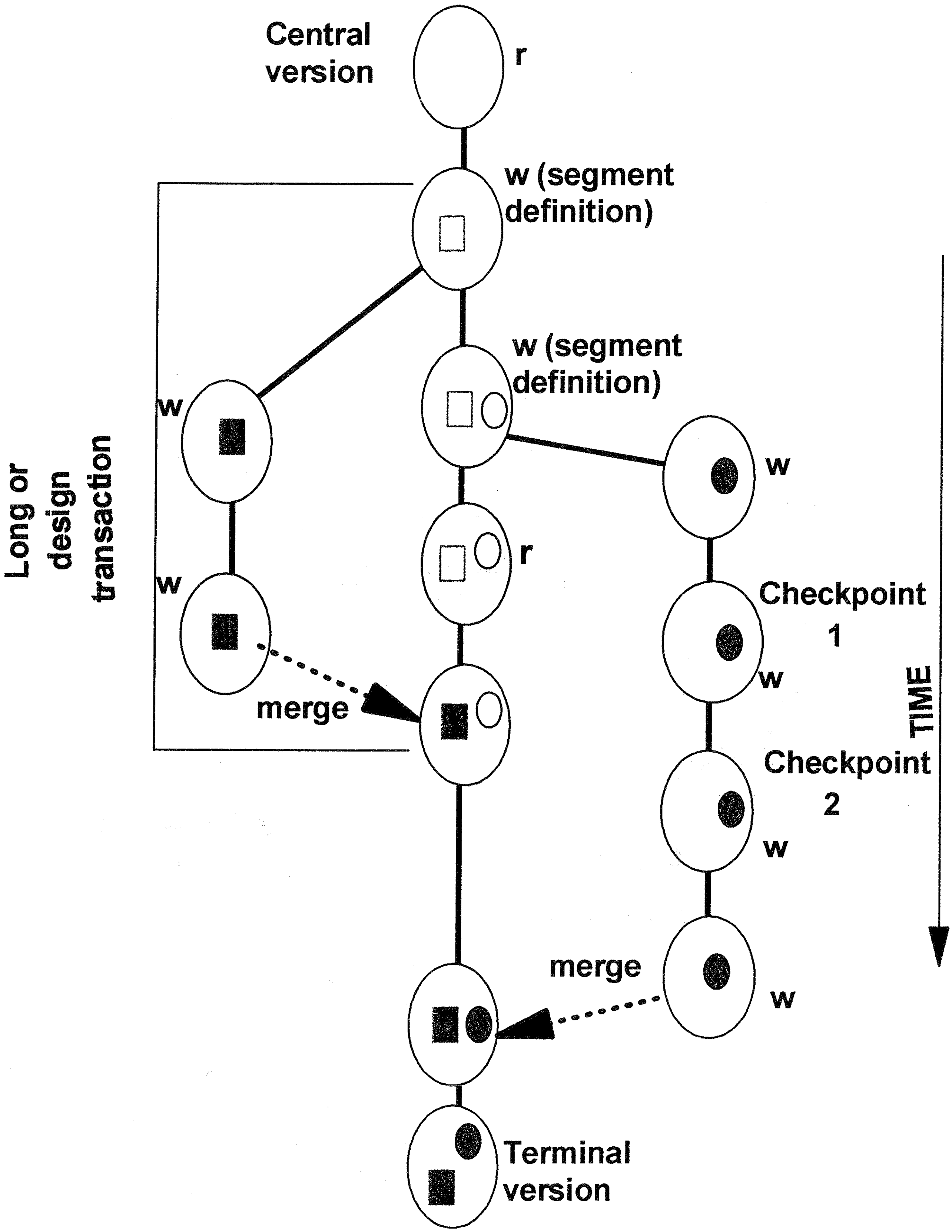

Figure 2: Versioning and Long Transaction Support

ACTIVITY	METHOD	REPRESENTATIVE FUNCTIONS
Create new object	CONSTRUCTOR	Issue unique obj-id
	VALIDATE	Correctness of attributes topology
Modify existing objects	PRE-MODIFY	Authority check
	POST-MODIFY	Internal checks
	VALIDATE	Checks in relation to other objects (geometry, topology)
		Management Log
		History Object?
Destroy object(s)	DESTRUCTOR	Authority check
	VALIDATE	Resolve trailing references
		History object?
Merge child version into parent	VALIDATE	Checks with respect to whole dataset
	MERGE	Management Log and generation of update notices and guidance files (digital updates)
	COMMIT [ROLL BACK]	Update metadata

Figure 3: Object Life Cycle