

The Evaluation of Educational Service Integration in Integrated Virtual Courses

Claus Pahl
School of Computer Applications
Dublin City University
Dublin 9
Ireland
cpahl@compapp.dcu.ie

Abstract

The effectiveness of an integrated virtual course is determined by factors such as the navigability of the system. We argue that in a virtual course, which offers different educational services for different learning activities, the integration of services is a good indicator for the effectiveness of a virtual course infrastructure. We develop a set of metrics to measure the degree of integration of a virtual course. We combine structural metrics and an analysis of the student usage of the system in order to measure integration.

1 *Integrated Virtual Courses*

Third-level undergraduate and postgraduate courses, especially in science and engineering, comprise usually various forms of teaching. Lectures typically form the backbone. Tutorials illustrate and exemplify the concepts and theories demonstrated in the lectures. Laboratory sessions allow the students to work, exercise, experiment, or explore topics they have learnt about in lectures and tutorials. We propose a similar separation of teaching forms for virtual Web-based courses. In virtual courses we shall call these forms *educational services*. Each service allows a student to interact with the system in a specific way. Each service is connected to a specific student *learning activity*. Attending a lecture, and thus comprehending concepts and theory, is different from acquiring practical skills in labs. Pedagogically, these activities should be connected and well integrated in a coherent environment. The educational activities should also be distinguishable in virtual environments. The difference in the student activities suggests implementing them through different services in a virtual course environment. A separation of services has also other advantages: the services might already exist in separation, or the course has to be developed by several people, or the course has also to be taught in a conventional form on campus.

In this paper we discuss conceptual and technological requirements for service integration in virtual courses. We address the evaluation and design of such course systems. We briefly relate to some experience we made delivering such a course. Our focus is on Web-based course systems using Internet technologies to implement the services. The literature [BL99] uses the term *Virtual Learning Environment VLE* for software systems that provide methods for online delivery of course material and communications functionality (email, newsgroups, video-conferencing). We consider a special class of VLEs here, which we will call *Integrated Virtual Courses*, abbreviated *IVC*. These are courses typically taught online at undergraduate or postgraduate level in third-level institutions. They offer typically several services. We ignore the communications issues here. An IVC is a Web site with specific page types and a specific architecture. Pages are classified into portal, menu, index or content pages. The architecture includes several central hubs (portal, index and menus) and peripheral content pages. Possibly longer paths represent causal dependencies among the individual content elements of the course.

Our objective is to validate and, if necessary, to improve the quality, in particular the *effectiveness* of a course system. We define effectiveness as the usability of the system. In an online course which coherently

combines various activities (or services), the effectiveness depends on the navigability in general, and the integration of the services in particular. The effectiveness shall be evaluated by *evaluating* the degree of *integration*, i.e. by analysing the navigation structures and evaluate how well integrated the services are. We will present *methods* and *metrics* in order to analyse the degree of integration of a course system. Running IVCs has shown us that a high degree of integration is necessary in order motivate students to use a system in a learning-oriented way. Coherence of the course content is certainly of paramount important for the success of a course. However, we will not look at course content authoring here, rather focus on the technical side of virtual course development. We will address common problems of Web-based systems such as disorientation and cognitive overhead. Coherence content should be reflected by the structure of the system. Looking at the classical delivery of courses on campus, we certainly expect the content of lectures, tutorials and labs to be coherent. For integrated virtual courses we need coherence also on the technical side, i.e. the course services need to be connected and co-ordinated. The key questions are how to measure integration and how to design integrated courses? We shall start with evaluation techniques and derive some design ideas from the evaluation criteria. The navigation structure of the virtual course can be measured using standard metrics, such as compactness or stratum, applied to the educational context. Besides the topology of the course Web site, the dynamic aspects such as the actual usage (the user behaviour) is the second important evaluation criterion. Essential for the evaluation is the combination of structural and behavioural criteria.

2 Evaluation of Integrated Virtual Courses

The effectiveness of a virtual course is determined by the usability and the degree to which a learning-oriented system has been implemented. Integration of services (or activities) is a property of an IVC, which expresses effectiveness in technical terms. We will show how integration can be measured and evaluated. Integration represents the structural coherence of the course system. The following criteria shall illustrate effectiveness and integration of an IVC:

- *Adequacy of page integration*: hubs like index pages should be central and accessible (a structural criterion), important content pages should have a high number of visits (a behavioural criterion). If the latter is not the case, a higher centrality/accessibility might be a remedy.
- *Navigability of the course structure*: compactness and the degree of choice are structural criteria, session logs are behavioural criteria. An erratic behaviour might result from a poorly structured site (indicated e.g. by too much choice or too little guidance). The existence of paths between related course elements and the ease with which these paths can be located are other issues.

Each of these criteria involves structural (static) and behavioural (dynamic) aspects. Thus, an evaluation method needs to combine structural and behavioural metrics. Both forms have been used to evaluate hypertext systems, but it will be shown here that a coherent integration of both will result in a better analysis of IVCs (and thus leading to improved IVC implementations). The basic information on which the evaluation will be carried out is the Web-site structure. The course structure, also called the *topology* of the course, consists of documents and links, i.e. forms a hypertext structure. Instead of simple graphs, the standard choice, we will use Petri-nets to model the topology of an IVC.

The evaluation of the effectiveness of a course system can be carried out based on the following methods:

- *Structural metrics*: The topology, i.e. the static structure of the site, is analysed. Are pages central, easy to access? Is the site compact? We adapt standard metrics for hypertext systems to the educational context for the structural analysis.
- *Behavioural metrics*: The student behaviour is analysed. How do the students navigate? Which pages do they visit? We use techniques such as server logs for the behavioural analysis.

<i>Goal I</i>	:	to evaluate the effectiveness of the IVC.
<i>Question I</i>	:	what is the effectiveness of an IVC?
<i>Metrics I</i>	:	integration.
<i>Goal II</i>	:	to improve the level of integration.
<i>Question II.1</i>	:	what is the influence of the topology on the level of integration?
<i>Metrics II.1</i>	:	compactness, stratum (degree of choice).
<i>Question II.2</i>	:	what do the user's navigation strategies tell us about the level of integration?
<i>Metrics II.2</i>	:	session logs, usage pattern.

We attempt a systematic evaluation approach. We follow the so-called *GQM-approach (Goal-Question-Metric)*. The entity to be evaluated is integration – see box above. Variables to be measured by metrics include the size of application, the topology, the compactness, the degree of choice in the site, or the user behaviour. Normally, the user’s opinion is included in any analysis. We concentrate here on structural and behavioural analysis, but we would like to emphasise here the importance of including the user’s opinion.

2.1 Topology

The *topology* of a Web-site is normally represented as a *directed graph*. In our case graphs do not provide the necessary means to describe our hyperlink structure. Graphs allow us to describe the transition from one document (a node) to another by following a link (an edge). Browser technology allows us to use link types that display a new window besides the current one. This cannot be modelled by simple graphs. *Petri-nets*, essentially graphs with two kinds of nodes (places and transitions), are a more suitable notation for the structural description and also the behavioural analysis of Web-sites (see Figures 1 and 2 next page). Their usefulness for evaluating hypertext structures has already been shown [BHK94]. We use *Condition/Event-nets* here. Places represent pages, transitions represent links. Tokens are placed in places and moved around through the net in order to analyse the dynamics. Only one token is allowed in each place, indicating that a user currently views a particular page. Events correspond to following a link. We introduce different *transition types* for Petri-nets. Structural links connect content pages with the table of contents or connect an index with content pages. Referential links are used for cross-referencing within the content pages. The examples in Fig. 1 and 2 use structural links only.

2.2 Structural Metrics

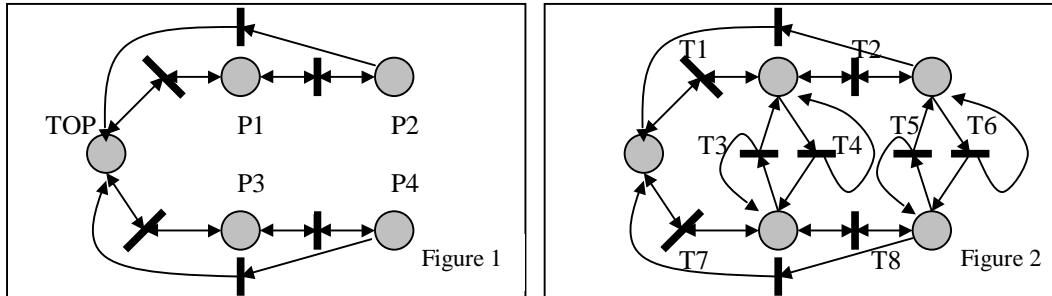
Based on a Petri-net we can derive a *distance matrix*, which allows us to adapt classical graph-based metrics to the educational context easily. An entry in the matrix is the length of the shortest path from a place A (a page) to a place B (another page) by counting places only. *Metrics* are derived from this matrix or a so-called converted distance matrix where infinite values (pages not reachable) are replaced by a constant. Metrics in general express ideas such as *compactness* of the site, or *centrality* and *accessibility* of nodes. They are a means to validate the technical realisation of the course structure. These metrics can be derived automatically, but need to be interpreted manually [BRS92].

<i>Metrics which apply to a single page are:</i>			
<u>Acronym</u>	<u>Name</u>	<u>Aspect</u>	<u>Definition</u>
<i>COD</i>	Converted Out Distance	Centrality	sum of all entries in a row
<i>CID</i>	Converted In Distance	Accessibility	sum of all entries in a column
<i>ROC</i>	Relative Out Centrality	Relative Centrality	CD/COD
<i>RIC</i>	Relative In Centrality	Relative Accessibility	CD/CID
<i>Metrics which apply to a whole site are:</i>			
<u>Acronym</u>	<u>Name</u>	<u>Definition</u>	
<i>CD</i>	Converted Distance	sum of all finite values in the matrix	
<i>CP</i>	Compactness	$(Max - CD) / (Max - Min)$ with $Max = (n^2 - n)C$, $Min = (n^2 - n)$, n : number of pages, C : maximal distance in matrix	

We can use metrics to classify pages in a course system. *Hubs* are central pages, i.e. have a high *ROC*, and are easy to find, i.e. have a high *RIC*. Examples are table of contents or index pages or any page part of the course portal. *Content pages* are less central, i.e. have a low *ROC*, with only a few links to other pages, i.e. a low *RIC*. The overall compactness of the site is one of the indicators for the degree of integration of a page. A high compactness indicates good navigability and good cross referencing, but also indicates a possibly poorly structured site.

The *stratum* metric reveals to what degree the hypertext is organised, for example that some nodes must be read before others. The *stratum* metric is based on the non-converted distance matrix. Metrics such as *OD* (Out Distance) and *ID* (In Distance) are defined as the sum of all finite entries. *D* is the total sum of all finite distances. The *prestige P* of a page is defined as $P = OD - ID$ for that page. The *absolute prestige AP* of the net is defined as the sum of all absolute values of prestiges of all pages. The *stratum S* is a

normalised absolute prestige, typically normalised by the *linear absolute prestige LAP* - the absolute prestige of a linear sequence of pages. The stratum of a web page indicates how much or little choice a user has while browsing a site. A high stratum means little, maybe too little choice. A low stratum indicates a lot of choice, but could also indicate an unstructured site without any clear dependencies between nodes. Instead of a linear sequence, we suggest a *normalisation by a reference educational site structure* as exemplified in Figure 1.



We apply the metric to a site consisting of two sequences of pages, (P1,P2) and (P3,P4), whose first elements are pointed at by a top node TOP and where each node has a link to TOP. This shall act as a minimal, or reference requirement for IVCs. If we add additional links which connect the i -th element in the first sequence with the i -th element of the second sequence (Fig.2), the stratum for simple sites results in the following values for n places in each sequence: $S=4/8$ ($n=2$), $S=16/24$ ($n=3$), and $S=40/50$ ($n=4$). Heuristics can now be established. Looking at larger numbers for n , stratum values around 0.8 seem to be the most suitable, indicating enough interrelations and a still structured site.

2.3 Behavioural Metrics

The individual *user behaviour* can be determined based on several questions. Which pages have been visited? In which order have pages been visited? In an educational system, other questions are also important. What is the total number of page visits (in relation to the number of students)? How often has a particular page been used? What is the distribution of visits over time (are there peaks)? Are navigation patterns emerging (wanted/unwanted/unexpected)? In particular the last point is important, since the answer to that question allows us to monitor and analyse how students learn, and whether they learn the way we expect them to learn. Intended and actual use (or intended and actual navigation pattern) can be compared. Students might learn differently with virtual course technology than we expect. Therefore, the integrated treatment of evaluation and design of such courses is of paramount importance.

Users in any hypertext system typically use one or a combination of *navigation strategies* [Voc94]. Breadth-first navigation visits first all nodes that are only one link away, i.e. read first a neighbouring node, then come back, then the next neighbouring node, and so on. Depth-first follows a given link sequence to go further and further away from the starting node. Given our simple topology model, breadth-first navigation corresponds to a vertical navigation, moving between different services covering the same problem. Depth-first compares to the conventional approach of following a sequence of lecture pages or lab exercise pages. Navigation in general comprises activities such as wayfinding (destination is known), exploration (not destination-guided), and object identification (location of destination not relevant). An example of a *navigation pattern* in an IVC is a lecture or lab participation with a mostly sequential visit of causally related pages by following structural links. In a revision phase the navigation is much more determined by the use of search and index facilities.

In order to evaluate user behaviour we define *sessions* based on the individual logs produced by the Web server. These sessions, representing dynamic behaviour, have to be mapped onto the static net topology. A session is a representation of the exact user behaviour. In order to analyse the course site we need to abstract various sessions into usage *pattern*, which can then be compared with patterns defined by the course developer. These patterns describe causal dependencies and the intended use, respectively. The edges describe the navigation, i.e. the order in which links are followed. Links, which are part of the prescribed pattern, should be emphasised or given more weight in the system implementation. *Path*

expressions describe the pattern: e.g. [T1 T4 T2 T6 | T7 T3 T8 T5] offers a choice between two sequences. The first sequence navigates via link T1 to page P1 (e.g. a lecture), then opens in parallel P2 (the corresponding lab) using link T4, then goes on to the next lecture page P2 using T2, and again opens the corresponding lab page P4 in parallel using T6. Path expressions allow us to specify choice S|T, parallel composition S||T, iteration T*, T⁺ and options [T]. An *ordering* S < T on path expressions compares actual and intended use and decides whether the actual use conforms with the intended use. T⁺ ≤ T means that actual repetitions are allowed, S ≤ S|T means that the user can choose between S and T, and SU ≤ S|T|U means that optional transitions can be left out. ≤ shall be reflexive, antisymmetric, and transitive, i.e. should form a partial ordering. A weaker variant shall also be introduced: STU ⊆ XY if S ≤ X and U ≤ Y which allows students to deviate for a while from the pattern. Deviation and repetition in actual navigation sequences are important patterns for the understanding of the way students work with the system. The relation of affected pages to their static structural metrics can improve the interpretation of these patterns.

3 Design of Integrated Virtual Courses

In this section we will briefly discuss design issues emerging from the evaluation techniques presented in the previous section. Relevant criteria for an integrated design include the compactness of the Web site or the proper positioning of content and hub pages. The basis for the development shall be formed by an *abstract model* for an *integrated educational service infrastructure*. The model should essentially prescribe a layered navigation topology. Each layer corresponds to one of the educational services. The pages in each layer are structurally linked according to their causal dependencies. Between the layers, links connect pages covering the same topic (but implementing different learning activities). Therefore, we can distinguish two forms of using a system based on our model:

- *Horizontal*: navigating within a layer, corresponds to depth-first navigation. This means using the infrastructure in a conventional way. The lecture service is used for a while, then the lab afterwards.
- *Vertical*: navigating between the layers, corresponds to breadth-first navigation. All services are used which provide different perspectives of the same problem. One problem is deeply worked on in theory and practice before the next one is addressed.

The course developer specifies navigation pattern based on these usage forms. Clearly, vertical use is problem-oriented and therefore more learning-oriented. In order to classify the degree to which vertical use is possible, we suggest *integration levels* for the layered topology. *Level0* means no integration. *Level1* means content integration, but no structural service integration. *Level2* classifies an integrated service topology. Here we can furthermore distinguish between a typically use in horizontal mode and use in vertical mode.

The infrastructure model needs to contain a meta-structure for the content part, which consists of portal, index and other menu pages. Disorientation, i.e. the problem of getting lost, can be countered through this meta-structure. The separation into different services also simplifies orientation (and also reduces cognitive overhead). Horizontal navigation is service-oriented and vertical navigation is problem-oriented. Cognitive overhead is a problem which affects the author: creating, naming, keeping track of content topics; and also the user: making decisions which links to follow, which links to abandon. Our layered model gives a basic structure reducing the problem.

4 Experience: an Integrated Virtual Database Course

CA309 is a virtual undergraduate Database course [Pah99]. It provides lectures, tutorials and labs online. The *lecture service* combines audio and visual material [SK99]. Due to the interactivity of the audio material we have obtained an individualised lecture service. The *tutorial service* provides a dynamic interactive illustration of the lecture material. A Web-based interface to a database server enables students to execute database queries, which are presented and explained in the lectures, in interaction with the system. The lab service is based on the same technology, but adds different examples, more complexity and new problems to be solved by the students. The system is not a mere reproduction of conventional course material. It has been designed for the specific medium. The result is an individualised interactive virtual course system. It falls into the same category as other systems, which use simulations in order to illustrate certain topics dynamically and interactively.

We omit details of the effectiveness evaluation here. We only summarise a few results and draw some conclusions, see [Pah00] for more details. Hardly any vertical use from lecture to lab occurred without appropriate navigation support, whereas an occasional vertical use from lab to lecture happened (looking up concepts, definitions, etc.). The system was mostly used in conventional horizontal mode: a lecture session, then a lab session. Lectures were used on a more regular basis than tutorials and labs. Our experience is that it is difficult to break the culture of consecutively attending lectures, tutorials, and labs without a well-designed and implemented navigation system.

5 Conclusions

Virtual courses provide different educational services. These services provide facilities for different learning activities. Different forms of interactivity have to be integrated. An integrated course system based on our layered model suggests using an IVC in different forms. In a vertical mode the students address one particular topic or problem at a time. In learning-oriented educational systems students work based on problems which connect theoretical and conceptual knowledge with the practical skills needed - students learn by doing. Theory is looked at when needed, the tutorials are used to illustrate the theory and put it into practice, and the labs are used to solve more related exercises within the same topic. This emphasises the close relationship of problem-orientation and integration. An integrated system allows problem-oriented learning. Besides the pedagogical advantages this approach has also a practical advantage: a mix of different activities is less monotonous and keeps the student's attention high. The objective of this paper was to relate structural and behavioural evaluation in order to obtain a better assessment of the effectiveness of IVCs. Metrics for both aspects complement each other.

In [MHH97] the application of metrics for authoring of educational hypertext systems is discussed. It is stated that the overall structuring of the knowledge domain (structural knowledge design) typically happens first before any representational planning in terms of documents and links takes place (topological design). A rich language for topological design providing e.g. various link types has proved to be important in order to make the knowledge structure as explicit as possible and give coherence and context to the users. Our infrastructure model is such a design tool. Metrics are developed in [MHH97] in order to support the maintainability of the system and the reusability of information, and to minimise the development effort. Our objective was to improve the effectiveness of IVCs, which we believe can only be guaranteed by an integrated system. Therefore, we have investigated static and dynamic metrics in order to evaluate the degree of service integration in an educational system, thus to validate the effectiveness of the system. In [MHH97] it is shown that most of the attributes also influence other properties of an educational system such as maintainability of the system or reusability of information.

Bibliography

- [BL99] S. Britain and O. Liber: *A Framework for Pedagogical Evaluation of Virtual Learning Environments*, Report JTAP Programme, 1999
- [BRS92] R.A. Botafogo, E. Revlin and B. Shneiderman: *Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics*. ACM Trans. on Information Systems 10(2), pp. 142-180, April 1992.
- [LH99] D. Lowe and W. Hall. *Hypermedia & the Web - an Engineering Approach*. John Wiley & Sons, 1999.
- [MHH97] E. Mendes, W. Hall and R. Harrison. *Applying Metrics to the Evaluation of Educational Hypermedia Applications*. Journal of Universal Computer Science. Vol. 4 (4), 1997
- [Pah99] C. Pahl. *CA309 Databases (Course Web Site), 1999/2000*, URL for Web site: <http://www.compapp.dcu.ie/~cpahl/teach/ca309/ca309.htm>.
- [Pah00] C. Pahl. *Delivery of a Virtual Integrated and Interactive Undergraduate Course*. Irish Educational Technology User's Conference, Sligo, Ireland, May, 2000
- [SK99] A.S. Smeaton and G. Keogh. *An Analysis of the Use of Virtual Delivery of Undergraduate Lectures*. Computers & Education, 32(1):83-94, 1999.
- [Voc94] J.W. de Vocht: *Experiments for the Characterisation of Hypertext-structures*. Master's Thesis. Eindhoven University of Technology, 1994