

Dynamic Presentation of Document Content for Rapid On-Line Skimming

Branimir Boguraev,¹ Christopher Kennedy,² Rachel Bellamy,³ Sascha Brawer,⁴ Yin Yin Wong,⁵ Jason Swartz⁶

Abstract

A wide range of current approaches to document summarisation utilise a strategy of data reduction over the original document source, whereby fragments of source text that are somehow identified as ‘most representative’ of the document’s content are extracted and presented to the user. In order for these fragments to be useful as summaries, it is necessary to know how they relate to the document; in order for these fragments to be usable as summaries, they must serve as windows into the document as a whole, with suitably designed interfaces for navigation into areas of particular interest. This paper discusses the notion of strong contextualisation of document highlights, how this translates into necessary features for document analysis, and how the document abstractions derived from such principles facilitate dynamic delivery of document content. We argue that dynamic document abstractions effectively mediate different levels of granularity of analysis, from terse document highlights to fully contextualised foci of particular interest. We describe a range of dynamic document viewers which embody novel presentation metaphors for document content delivery.

1 Introduction

Present day summarisation technologies fall short of delivering fully informative summaries of documents. Largely, this is due to shortcomings of the state-of-the-art in natural language processing; in general, the issue of how to customise a summarisation procedure for a specific information seeking task is still an open one. However, given the rapidly growing volume of document-based information on-line, the need for *any* kind of document abstraction mechanism is so great that summarisation technologies are beginning to be deployed in real world situations.

The majority of techniques for ‘summarisation’, as applied to average-length documents, fall within two broad categories. One type of approach involves mining a document for certain pre-specified pieces of information, typically defined *a priori* as characteristic features of a known domain of interest; this information is then used to generate an abstraction of the document’s content. An alternative type of analysis seeks to ‘re-use’ certain fragments of the original text that have been somehow identified as representative of the content of the document as a whole (or of some coherent segment of the document). A variety of approaches fall into this general category, ranging from fairly common sentence extraction techniques to newer methods utilising, for example, strong notions of topicality [4], [8], lexical chains [3], and discourse structure [14], [5] (see the papers from the recent ACL workshop on Intelligent, Scalable Text Summarization [2] for relevant overview). Ultimately, all of these approaches share a fundamental similarity: they construct a characterisation of document content through significant reduction of the original document source, rather than through some kind of generation procedure. This raises several important questions.

First and foremost, what is the optimal way of incorporating the set of extracted, topically indicative fragments into a coherent representation of document content? Unlike techniques which rely on domain models—or other ‘strong’ knowledge sources—to generate summaries, data reduction

approaches typically pay little attention to the problem of synthesizing a coherent summary explicitly tailored for clarity and cohesion of the information. Included in this question is an issue of granularity of data reduction: what sorts of expressions make the best information-bearing passages for the purpose of summarisation? Are sentences better than paragraphs? Are phrases even better?

The second question involves the level of user involvement. From the end-user’s point of view, making judgements about a document on the basis of a summary involves a sequence of actions: look at the summary, absorb its semantic impact, infer what the document might be about, decide whether to consult the source, somehow call up the full document, and navigate to the point(s) of interest. How can a summary, then, alleviate the cognitive load placed on a user faced with a large, and growing, number of documents on a daily basis?

Finally, acknowledging that different information management tasks may require different kinds of summary, even from the same document—a point made recently by Sparck Jones [16], how should the data discarded by the reduction process be retained, in case a reference to a part of the document not originally included in the summary is required?

This paper opens a discussion of these questions, and offers some initial answers for them. In particular, we argue that in order for summaries derived by extraction techniques to be useful, they must satisfy two constraints: they must incorporate a granularity of reduction down to *phrasal* analysis, and they must be presented to users through *dynamic* interfaces. We demonstrate that such a summarisation technology facilitates a process of ‘filling in the gaps’ left by reduction of the source and leads the user deeper into the content of the original document, while retaining strong notions of contextualisation as an inherent property of the discourse. The organization of the paper is as follows.

In Section 2, we analyse certain usability aspects of summarisation technologies, and argue for a range of features of the analysis of a document which need to be retained as an

¹Advanced Technologies Group, Apple Computer, Inc., Cupertino, CA 95014, bkb@cs.brandeis.edu

²Department of Linguistics, Northwestern University, Evanston, IL 60208, kennedy@ling.nwu.edu

³Advanced Technologies Group, Apple Computer, Inc., Cupertino, CA 95014, rachel@acm.org

⁴Department of Computer Science, University of Zürich, Switzerland, brawer@coli.uni-sb.de

⁵Wong Design, San Francisco, CA 94111, yinyin@media.mit.edu

⁶Advanced Technologies Group, Apple Computer, Inc., Cupertino, CA 95014, jswartz@sinewave.com

integral part of any abstraction or summary for that document, including contextualisation of document highlights. We then sketch an interface environment for delivering such abstractions to end-users, in which a strong notion of context is maintained throughout the interaction between users and documents by dynamic delivery of document content. In Section 3, we outline a technology for phrasal-based summarisation, and in Section 4, we discuss a range of experiments with dynamic visualisations of document content, introducing temporal typography as a particularly promising vehicle for dynamic document delivery. Section 5 describes a range of dynamic document viewers which implement novel modes of summary presentation and content visualisation.

2 Document summarisation in use

2.1 Behind document ‘summaries’: Gaps in understanding

To the extent that broad coverage summarisation techniques are beginning to get deployed in real world situations, it is still the case that these techniques are based primarily on sentence extraction methods. To illustrate some of the issues raised by the questions listed above, consider an example of such a technique, taken from an operational news tracking site.⁷ Under the heading of *Articles about IRS Abuses Alleged*, some entries read:

RENO ON SUNDAY / Reform Taxes the...

The problem, of course, is that the enemies of the present system are all grinding different axes. How true, how true, and ditto for most of the people who sit on the Finance Committee. (First found: 18 Oct 1997)

Scheduled IRS Layoffs For 500 Are...

The agency’s original plan called for eliminating as many as 5,000 jobs in field offices and at the Washington headquarters. “The way this has turned out, it works to the agency’s advantage, the employees’ advantage and the union’s advantage. (First found: 17 Oct 1997)

Both examples present summaries as sentences which almost seamlessly follow one another. While this may account for acceptable readability, it is at best misleading, as in the original documents these sentences are several paragraphs apart. This makes it hard to know, for example, that the expression “*How true, how true*”, in the first example is not connected to anything in the first sentence, but rather to some expression elsewhere in the ‘invisible’ portion document. Similarly, the thrust of the second article—namely that there is a reversal of an anticipated situation—is not at all captured: it turns out that the missing paragraphs between the summary sentences discuss how the planned 5,000 layoffs have been reduced to “4,000, then 1,400 and finally settled at about 500”, and that “now, even those 500 workers will not be cut”. Some indication to this effect might have been surmised from the full title of the article, *Scheduled IRS Layoffs For 500 Are Cancelled*; unfortunately, this has been truncated by a data reduction strategy which is insensitive to notions of linguistic phrases, auxiliary verb constructions, mood, and so forth.

In the extreme case, such summaries can range from under-informative (as illustrated by the first example above), to misleading (the second example), to plainly devoid of any useful information, as in the case of a third example from the same site:

Technology News from Wired News

This is more than 500 times thinner than a human hair. ‘Don’t expect one in a present under your Christmas tree this year’.”

2.2 Filling in the gaps (1): Granularity of analysis and phrasal identification

This discussion is not intended to highlight the drawbacks of a particular summarisation technology. The purpose is to show how hard it is to fill the gaps that are necessarily introduced by an approach to summarisation that selects certain fragments from the original source as representative of document content. Presently, the only way of filling such gaps is by the user actively requesting the entire document, a time- and attention-consuming procedure. An alternative strategy would incorporate into the document analysis technology principles that would facilitate the process of filling in the gaps without full document retrieval. Two important principles in this regard are the related notions of *granularity of document analysis* and *the importance of linguistic phrases as units of representation*.

Granularity is closely tied to context. In general, the optimal way to make sense of the information conveyed by a given sentence can be augmented by positioning it in its paragraph context; similarly, the theme and topic(s) in a paragraph can be further elaborated by relating it to the segment of discourse which encompasses the theme in its entirety. This view of context provides a natural containment hierarchy, relating the different information levels in a document together. Such a hierarchy also extends in sub-sentential direction: phrasal units indicative of topicality are clearly wholly contained in sentences; furthermore, a phrasal containment hierarchy could also be utilised to provide progressively more detailed information about the topical phrases in their context.

We will refer to this process below as *contextualisation of topical phrases*. Before we move on, however, two points of clarification should be made. First, the levels of granularity of contextualisation described here, which range from phrases at the most focused view, through clauses, sentences and segments of discourse, to the entire text at the broadest level, have been adopted on the basis of the technology we are working with (built around a phrasal grammar, as discussed in Section 3), not from larger theoretical concerns. In principle, other levels of granularity could prove to be equally appropriate for the task at hand. Second, our use of the terms ‘context’ and ‘contextualisation’ should be distinguished from other uses, which might refer to, for example, the set of propositions assumed to be true at a particular point in the discourse or the set of entities that have been introduced by (possibly coreferential) linguistic expressions in the text. Although ‘context’ in this latter sense clearly plays a fundamental role in determining the information content of a particular text (indeed,

⁷The *News Channel* page of Excite, an information vendor and a popular search engine host for the World Wide Web, is available via the “Ongoing Coverage” section of the news tracking page, <http://nt.excite.com>.

it underlies our core summarisation algorithm; see Kennedy and Boguraev [12] and Boguraev and Kennedy [4] for detailed discussion), our focus here is on how a granular view of ‘textual context’ can be used as the basis for an informative and useful *representation* of document content.

For an illustration of our notion of contextualisation, imagine that in the second example above some mechanism has determined that the phrase “*Scheduled IRS Layoffs*” is a topic (see 3.2 below for more detailed discussion of topicality and the notion of topical phrases). Assuming some focused mining in the vicinity of such an ‘anchor’ by a phrasal grammar of a certain type, this topical phrase could be further contextualised to “*Scheduled IRS Layoffs For 500 Are Cancelled*”. Similar expansion of topic in context might yield, for the initial discourse segment of the document, progressively larger and more informative fragments from it:

❑ TOPIC STAMP:	<i>“Scheduled IRS Layoffs”</i>
❑ TOPIC IN RELATIONAL CONTEXT:	<i>“there will be no layoffs”</i>
❑ TOPICAL SENTENCE:	<i>“Yesterday, the IRS said there will be no layoffs”</i>
❑ SENTENCE IN PARAGRAPH CONTEXT:	<i>“More than a year ago, the Internal Revenue Service planned widespread job cuts. Yesterday, the IRS said there will be no layoffs.”</i>
❑ PARAGRAPH WITHIN TOPICALLY COHERENT DISCOURSE THEME:	<i>“More than a year ago, the Internal Revenue Service planned widespread job cuts. Yesterday, the IRS said there will be no layoffs. Confronted with congressional criticism and calls for reform in light of some highly publicized reports of abusive actions toward taxpayers, as well as staunch union opposition to the cuts, the IRS said employees at risk of losing their jobs would be reassigned to improve ‘customer service,’ help taxpayers resolve problems and increase compliance with tax laws.”</i>

The containment hierarchy of layered information in this example—from very compact and representative topical phrases all the way to full and rich discourse segments—represents and maintains a strong notion of contextualisation in the document abstraction. In order for this approach to work, however, it is crucially necessary to identify phrasal units smaller than sentences—i.e., noun phrases, verb phrases, and clausal units—and to arrange them in layers corresponding to the informational containment hierarchy. Without phrasal analysis and a phrasal foundation to document abstraction, the multiple levels of granularity represented in this example cannot be constructed.⁸

2.3 Filling in the gaps (2): User involvement

It is clear that granularity of analysis and containment hierarchy of information-bearing phrasal units can be utilised very effectively to implement a ‘zooming’ function into and/or out of a given document. In this way finding out more of what is behind a document ‘summary’ is, in effect, filling in the gaps in such a summary in a controlled fashion, guided by incrementally revealing progressively larger and more informative contexts.

Conceptually, this is not dissimilar to the notion of ‘percentage of shrink factor’, typically utilised by sentence-based summarisers, where a user can specify that a document

should be condensed to $N\%$ of its full extent. There is, however, a crucial difference here. When re-casting a document from, say, 10% to 20% shrink factor, there is no way to specify ahead of time, nor to know after the event, how the additional sentences relate to the original 10%. In contrast, when a document is re-cast in terms of information-bearing units a level higher than what its current representation uses—for instance, as a set of relational contexts immediately surrounding its topical phrases—there is a guarantee that the user’s sense of what the document is about is monotonically enriched.

This approach puts the process of filling the gaps in summaries under more direct user control. Most current approaches employ a relatively rigid mechanism (typically sensitive to a mouse click, or some similar interactive command) with the simple semantics of “bring up the entire document, possibly with the point of view focused on the particular sentence of the summary which received the click, presented in its natural document context, and maybe highlighted”. Clearly, having a richer data structure—such as a containment hierarchy of information-bearing units—facilitates greater flexibility in interactions with what would be, in effect, a whole range of dynamically reconfigured summaries at different level of granularity and detail.

There is still one problem, however: the process of filling in the gaps requires active user involvement. In principle there is nothing wrong with this. In practice, real information management environments involve working with a large number of documents. It is far from clear that users will have the energy, time, dedication, and concentration required to assess, absorb, and act upon summaries for each one of these documents, by clicking their way through each member of a long static list. As a result, the mode of presentation of summary and the nature of the interface through which this presentation is done becomes crucial.

2.4 Filling in the gaps (3): Dynamic delivery

Our solution to the problem of effectively communicating to the end user the ‘gist’ of an on-line document, or of a collection of on-line documents, is based on the idea of relating form and content, by means of dynamic visual treatment of written language, or *temporal typography* ([17]). Only recently we have begun to appreciate, and experiment with, the possibility of escaping the static and rigid constraints of writing on paper.

Imagine you are looking at a small area on a computer screen. Words appear and disappear on the screen one by one. As they appear, meaning is expressed as forms change dynamically over time. The combined effect of the message, form and rhythm express a tone of voice, emotion or personality as if you hear a person speak. Although the two mediums, spoken and written words, are vastly different, the analogy may give you a sense of the expressive potential of temporal typography. (Wong, [17])

The notion, essentially, is to relate highlights of the core meaning of a message to ways of visually enhancing their impact,

⁸In addition, the identification of phrases supports the application of certain semantic operations crucial to the construction of a larger picture of contextual links across the document, such as reference identification, co-referentiality, and topic tracking; examples here would be relating “*layoffs*” to “*scheduled IRS layoffs*”, identifying “*Internal Revenue Service*” and “*IRS*” as referring to the same object, resolving anaphora in general, and so forth. See [12], [4] for relevant discussion; see also Section 3.

or at least mimicking (some of) their semantic load. In the immediate context of this discussion, this translates to questions of what might be appropriate visual metaphors for representing semantic objects like topical phrases, shifts in discourse structure, or contextualisation of information-bearing phrasal units, at different levels of granularity, as described in the previous section.

There are several appealing aspects to dynamically presenting abstractions of document content. The user need not be actively involved: as documents arrive at the desktop, they can be analysed and the resulting content abstractions can be displayed autonomously. Should the user have the time or inclination to focus on a particular document, interactive controls will be at their disposal; alternatively, each new arrival can be presented under its own schedule, followed by another, and so on. The presentation cycle can be customised to make use of arbitrary combinations of granularity of expressiveness. Notions like semantic highlights and demarcation of context are easily mapped onto visual metaphors, and thus naturally support the expression of content by means of variations of form. Cognitively, short phrases with high semantic load are amenable to punctuated display following a natural rhythm of visual perception.

In summary, delivering document content abstractions dynamically makes it possible to fully exploit the variable depth analysis of documents that we argue for here, enables the handling of continuous flow of information into one's personal workspace (see Section 4.1), and allows for smooth integration of passive absorption of the analyses by the end-user with active participation in more focused document perusal.

The following section briefly presents an analysis technology which seeks to derive document content characterisations designed to exhibit the semantic properties described above; in Section 4, we return to a description of some essential features of temporal typography, as it relates to dynamic delivery of document content.

3 Capsule overviews of documents

Elsewhere ([12], [4]) we have described a document analysis and content characterisation technology which develops and implements the ideas highlighted in Section 2.2. Striving to balance the conflicting requirements of depth and accuracy of a summary with those of domain- and genre-independence, we develop the notion of *capsule overviews* as content abstractions for text documents explicitly designed to capture 'aboutness' ([4]). This is represented as a set of *highly salient*, and by that token most representative, phrases in the document. Viewing topicality in its stricter, linguistic, sense, we define *topic stamps* to be the most prominent of these phrases, introduced into, and then elaborated upon, the document body. On the basis of this definition, we have developed a computational, algorithmic, procedure for generating a set of abstractions for the core meaning in the document, ultimately resulting in a *capsule overview* of the document based upon suitable presentation of the most representative, and most contentful, expressions in the text. These abstractions comprise layered and inter-related phrasal units at different levels of granularity, following the containment hierarchy presented in Section 2.2.

3.1 Document characterisation by topics

Topic stamps are certain noun phrases judged to be topically relevant in the text. One of the functions of a noun phrase in telling a story is to introduce a new entity—a *discourse referent*, typically an object, a concept, an event, a participant in an action—into the story. *Topical relevance* is defined as a feature of a discourse referent, which marks it as being subsequently elaborated upon in the course of story-telling. Following processes of identification and normalisation, discourse referents are ranked according to a global measure of *salience*. Salience of a topic is defined as a single numeric parameter, which embodies a number of semantic criteria. Some of these are: how prominently the topic is introduced into the discourse; how much discussion is there concerning the topic; how much is the topic mentioned throughout the entire document, as opposed to e.g. in just one (or some) sections of the text. The full set of semantic factors embodied in the salience weight is described in detail in [11] and [12]). The intent is to be sensitive to a number of linguistic and stylistic devices employed in text-based discourse for the purposes of introducing, defining, refining, and re-introducing discourse referents. The set of such devices is large, and it is precisely this richness which enables finer distinctions concerning content elaboration to be observed and recorded.

Encoding the results of such analysis into a single parameter provides the basis of a decision procedure which can extract the topics with high salience weight. These are the *topic stamps* for the document. While simple, the decision procedure is still remarkably well informed, as the salience weight calculation by design takes into account the wide range of ways in which topicality manifests itself in written prose.

The final set of topic stamps is representative of the core document content. It is *compact*, as it is a significantly cut-down version of the full list of document topics. It is *informative*, as the topics in it are the most prominent discourse referents in the document. It is *representative* of the whole document: in breadth, as a separate topic tracking module effectively maintains a record of where and how discourse referents occur in the entire span of the text, and in depth, as each topic stamp maintains its relational and larger discourse contexts. As topics are the primary content-bearing entities in a document, the topic stamps offer *accurate* approximation of what that document is about.

3.2 Topic stamps and capsule overviews

Capsule overviews of documents take the form of a set of topic stamps, enriched by the textual contexts in which they are encountered in the source. The topic stamps are organised in order of appearance, and are 'superimposed' onto progressively more refined and more detailed discourse fragments: relational contexts, sentences, paragraphs, and ultimately discourse segments (see 2.2 above).

Discourse segments reflect (dis-)continuity of narrative and the way in which focus of attention/description changes with the progress of the text story. 'Chunking' the document into more manageable units is not just for convenience. Discourse segments correspond to topically coherent, contiguous sections of text. The approach to segmentation we adopt implements a similarity-based algorithm along the lines of the one developed by Hearst [6], which detects changes in topic

by using a lexical similarity measure. By calculating the discourse salience of referents with respect to the results of discourse segmentation, each segment can be associated with a listing of those expressions that are most salient within the segment, i.e., each segment can be assigned a set of topic stamps. The result of these calculations, a set of segment-topic stamp pairs ordered according to linear sequencing of the segments in the text, is returned as the capsule overview for the entire document.

3.3 Capsule overview: an example

The operational components of salience-based content characterisation fall in the following categories: discourse segmentation; phrasal analysis (of nominal expressions and relations); anaphora resolution and generation of a referent set; calculation of discourse salience and identification of topic stamps; and enriching topic stamps with information about relational context(s). Some of the functionality derives from phrasal identification, suitably augmented with mechanisms for maintaining phrase containment; in particular, both relation identification and extended phrasal analysis are carried out by running a phrasal grammar over a stream of text tokens tagged for morphological, syntactic, and grammatical function (this is in addition to a grammar mining for terms and, generally, referents). Base level linguistic analysis is provided by a supertagger, [10]. The later, more semantically-intensive algorithms are described in detail in [11] and [12].

"ONE DAY, everything Bill Gates has sold you up to now, whether it's Windows 95 or Windows 97, will become obsolete," declares Gilbert Amelio, the boss at Apple Computer. "Gates is vulnerable at that point. And we want to make sure we're ready to come forward with a superior answer."

Bill Gates vulnerable? Apple would swoop in and take Microsoft's customers? Ridiculous! Impossible! In the last fiscal year, Apple lost \$816 million; Microsoft made \$2.2 billion. Microsoft has a market value thirty times that of Apple.

Outlandish and grandiose as Amelio's idea sounds, it makes sense for Apple to think in such big, bold terms. Apple is in a position where standing pat almost certainly means slow death.

It's a bit like a patient with a probably terminal disease deciding to take a chance on an untested but promising new drug. A bold strategy is the least risky strategy. As things stand, customers and outside software developers alike are deserting the company. Apple needs something dramatic to persuade them to stay aboard. A radical redesign of the desktop computer might do the trick. If they think the redesign has merit, they may feel compelled to get on the bandwagon lest it leave them behind.

Lots of "ifs," but you can't accuse Amelio of lacking vision. Today's desktop machines, he says, are ill-equipped to handle the coming power of the Internet. Tomorrow's machines must accommodate rivers of data, multimedia and multitasking (juggling several tasks simultaneously).

We're past the point of upgrading, he says. Time to scrap your operating system and start over. The operating system is the software that controls how your computer's parts (memory, disk drives, screen) interact with applications like games and Web browsers. Once you've done that, buy new applications to go with the reengineered operating system.

Amelio, 53, brings a lot of credibility to this task. His resume includes both a rescue of National Semiconductor from near-bankruptcy and 16 patents, including one for convening the charge-coupled device.

But where is Amelio going to get this new operating system? From Be, Inc., in Menlo Park, Calif., a half-hour's drive from Apple's Cupertino headquarters, a hot little company founded by ex-Apple visionary Jean-Louis Gasse. Its BeOS, now undergoing clinical trials, is that radical redesign in operating systems that Amelio is talking about. Married to hardware from Apple and Apple clones, the BeOS just might be a credible competitor to Microsoft's Windows, which runs on IBM-compatible hardware.

We illustrate the procedure by highlighting certain aspects of a capsule overview of a recent *Forbes* article [9]. The document focuses on the strategy of Gilbert Amelio (former CEO of Apple Computer) concerning a new operating system for the Macintosh. Too long to quote in full (approximately four pages in print), the passage from the beginning of the article shown here contains the first three segments, as identified by the discourse segmentation component; segment boundaries

are marked by extra vertical space (of course, this demarcation does not exist in the source, and is introduced here for illustrative purposes only).

The relevant sections of the overview (for the three segments of the passage quoted) are shown below. The listing of topic stamps in their relational contexts provides the core data for the capsule overview; while not explicitly shown here, the capsule overview data structure fully maintains the layering of information implicit in the containment hierarchy.

- 1: APPLE; MICROSOFT
APPLE would swoop in and take MICROSOFT's customers?
APPLE lost \$816 million;
MICROSOFT made \$2.2 billion.
MICROSOFT has a market value thirty times that of APPLE
it makes sense for APPLE
APPLE is in a position
APPLE needs something dramatic
- 2: DESKTOP MACHINES; OPERATING SYSTEM
Today's DESKTOP MACHINES, he [Gilbert Amelio] says
Tomorrow's MACHINES must accommodate rivers of data
Time to scrap your OPERATING SYSTEM and start over
The OPERATING SYSTEM is the software that controls
to go with the REENGINEERED OPERATING SYSTEM
- 3: GILBERT AMELIO; NEW OPERATING SYSTEM
AMELIO, 53, brings a lot of credibility to this task
His [Gilbert Amelio] resumé includes
where is AMELIO going to get this NEW OPERATING SYSTEM?
radical redesign in OPERATING SYSTEMS that AMELIO is talking about

The division of this passage into segments, and the segment-based assignment of topic stamps, exemplifies a capsule overview's 'tracking' of the underlying coherence of a story. The discourse segmentation component is sensitive to shifts in topic⁹—in this example, the shift from discussing the relation between Apple and Microsoft to some remarks on the future of desktop computing to a summary of Amelio's background and plans for Apple's operating system. Layered on top of segmentation are the topic stamps themselves, in their relational contexts, at a phrasal level of granularity.

The first segment sets up the discussion by positioning Apple opposite Microsoft in the marketplace and focusing on their major products, the operating systems. The topic stamps identified for this segment, APPLE and MICROSOFT, together with their local contexts, are both indicative of the introductory character of the opening paragraphs and highly representative of the gist of the first segment. Note that the apparent uninformativeness of some relational contexts, for example, '*... APPLE is in a position ...*', does not pose a serious problem. An adjustment of the granularity—at capsule overview presentation time (see below)—reveals the larger sentential context in which the topic stamp occurs, which in turn inherits the high topicality ranking of its anchor: '*APPLE is in a position where standing pat almost certainly means slow death.*'

For the second segment of the sample, OPERATING SYSTEM and DESKTOP MACHINES have been identified as representative. The set of topic stamps and contexts illustrated provides an encapsulated snapshot of the segment, which introduces Amelio's views on coming challenges for desktop machines and the general concept of an operating system. Again, even if some of these appear under-specified, more detail is easily available by a change in granularity, which reveals the definitional nature of the even larger context '*The OPERATING SYSTEM is the software that controls how your computer's parts...*'

The third segment of the passage is associated with the stamps GILBERT AMELIO and NEW OPERATING SYSTEM. The linguistic rationale for the selection of these particular noun phrases as topical is closely tied to form and function of discourse referents in context. Accordingly, the computational

⁹To the extent that such shifts are reflected in the lexical make-up of contiguous passages of the document: see [6], also see Section 3.2.

justification for the choices lies in the extremely high values of salience, resulting from taking into account a number of factors: co-referentiality between ‘Amelio’ and ‘Gilbert Amelio’, co-referentiality between ‘Amelio’ and ‘His’, syntactic prominence of ‘Amelio’ (as a subject) promoting topical status higher than for instance ‘Apple’ (which appears in adjunct positions), high overall frequency (four, counting the anaphor, as opposed to three for ‘Apple’—even if the two get the same number of text occurrences in the segment)—and boost in global salience measures, due to ‘priming’ effects of both referents for ‘Gilbert Amelio’ and ‘operating system’ in the prior discourse of the two preceding segments.

4 Temporal typography

Dynamic content delivery is based on ideas of temporal typography developed by Wong ([17]). This work is based upon synergy of psychological studies of reading, graphic design, and temporal presentation of text. Graphic design history is rich with examples of experimenting with visual treatment of written language. Designers have begun to explore temporal presentation of text in television and film media. Studies of reading, which to a large extent form the basis of Wong’s work, have explored dynamic presentation of content, related to the interactions between meaning and intent of a text-based message. However, Wong’s studies of the dynamic relationship between meaning and delivery formats assume that the annotations for meaning in her experiments have been done by hand. In contrast, here we focus on leveraging automatic document content analysis technology, capable of delivering meaning analyses and content abstractions precisely of the kind which can be effectively coupled with dynamic content delivery.

4.1 Visualisation of document content

In an earlier section (2.3), we briefly discussed the predominant current mechanism for mediating the spectrum between a summary of a document and a complete version of the same document. In addition to a direct hypertext rendering of extracted sentences, in their full document contexts, two variations on this approach are the VESPA slider and HYPERGEN. VESPA is an experimental interface to Apple’s sentence-based summariser ([1]), whose main feature is a slider which dynamically readjusts the shrink factor of a document summary. HYPERGEN exploits notions of phrasal containment within sentence units, in an attempt to elaborate the notion of granularity of analysis and context introduced in 2.2 above: in a process called sentence simplification, Mahesh ([13]) uses phrases as ‘sentence surrogates’, which are then straightforwardly rendered as hypertext links to the sentences themselves.

As part of an on-going investigation of visualizing large information spaces, researchers at Xerox PARC have looked at a variety of structured data types (such as hierarchically structured data, calendars, and bibliographic databases). Some general principles derived from that work have been applied to unstructured documents: the DOCUMENT LENS is a technique for viewing 2-D information, designed for component

presentations of multi-page documents. Without going into detail, what is of particular relevance here is the strong notion of *focus plus context* which drives the design. The visualisation, however, does little in terms of using any kind of document summary or other abstraction, and is of a predominantly static nature (even though it is extremely responsive to user interaction, as it attempts to combine a ‘bird’s eye view’ of the entire document with a page browsing metaphor). More recently, experimental prototypes have been developed for interfaces which treat term sets (in the information retrieval sense, i.e. flat lists of index terms) as document surrogates: the focus of such designs is on visually presenting the distribution of terms across the document, and on mediating access to local context for a given term ([15], [7]). Ultimately, however, these interfaces still offer only a direct link between two states, the document surrogate and its full form.

With the migration of news delivery over the World Wide Web and the growth of information ‘push’ vendors, we are beginning to see some methods for presentation of news stories which use notions of dynamic delivery of content. Most of these are variations on the same theme: news delivery using the ticker metaphor. Thus both ABC’s news site (<http://www.abc.com>) and Pointcast (<http://www.pointcast.com>) employ a traditional horizontal ticker, CNN Interactive (<http://www.cnn.com>) arrange their ticker vertically, CBS (<http://www.uttm.com>) combine a ticker with photos from a major story.

The important insight here is that tickers are dynamic objects, which can be programmed to continuously update themselves from a news feed and to cycle in a pre-defined regime, therefore not requiring user intervention (see 2.3 above). Furthermore, they can be dispatched to an area of the workspace (monitor screen) where constant, yet inobtrusive, news delivery can take place in the periphery of the user’s main activity: thus a choice exists between pro-active engagement with the news source, and passive (and almost subliminal) monitoring of news data.

None of the examples above, however, combines a ticker with an automatic summarisation engine. To a large extent this is because sentences—especially inconsecutive ones, in the absence of visual markers for discontinuity—do not lend themselves easily into the word by word, left to right, presentation mode. This is clearly a situation where phrasal units of a sub-sentence granularity can be used much more effectively. In addition, psychological experiments on active reading (see [17]) show that when text is presented dynamically in the manner of a ticker, subjects’ reading speeds are significantly slower than for text presented statically. On the other hand, dynamic presentations of text which show words or short phrases in the same location but serially, one after the other, have reading speeds comparable to those for normal static texts.

To date, no applications have been developed utilising temporal typography for dynamic delivery of content abstractions. Wong has looked at how dynamic type in general can be used for four different communicative goals: expressive messages, dialogue, active reading and real time conversation (see [17]). Most relevant to this discussion are her experiments on active reading. In one of these she used a basic RSVP¹⁰ method (words or phrases presented sequentially one after

¹⁰Rapid Serial Visual Presentation

another, on the same line and at the same position) to deliver a sequence of news headlines. In a second set of experiments called HIGHWAY NEWS, she used three dimensions, combined with a zooming motion, to present a sequence of text highlights. “News headlines are placed one after another in the z-dimension. Headlines are presented serially according to active input from the reader. The reader presses a mouse button to fly through the rows of headlines— as if flying over a highway of text.” ([17]). These experiments show the strong feasibility of high impact, low engagement, delivery of semantically prominent text fragments being utilised as a powerful technique for visualising certain types of inherently linear information.

None of the work cited above relies on automatically generated meaning abstractions as its input; yet, it is clear that the topically-rich capsule overviews generated by the document analysis technology discussed in Section 3 are just the kind of semantic highlights which Wong’s experiments in active reading assume. Conversely, up till now there has been no thought as to how the nature of topic-based capsule overviews in particular would fit the notion of dynamic type. This is the subject of the following section.

5 Dynamic document viewers

At the time of writing, three different viewers explore the ideas discussed above. The difference is largely due to the variety of operational environments in which the viewers have been applied. A variation on a news ticker is designed to be deployed in situations where screen real estate may be at premium, or where several different channels of information may be delivered simultaneously to the same ‘in-box’; typically such situations assume that users would only want to get a very general idea of document content. For situations where more screen real estate might be available, and/or it may be known ahead of time that more detail concerning document content might be required, a different viewer develops ideas from rapid serial visual presentation (RSVP). Yet another interface caters to the need to be able to get immediate access to the full text of a document, without losing the benefits of rapid skimming through content highlights while fully maintaining information about the larger context.

All of the viewers assume an environment where incoming documents get analysed to capsule overview level; the results of the analysis are embedded into the original text by means of special purpose tags.

5.1 TopicsTicker

TOPICSTICKER is a minimalist, hands-free, peripheral-vision-directed ticker tape, with certain aspects of its display tuned for serial delivery of a document’s topic stamps: the string in the left panel is the document title, and the right panel is where the display cycles through the document’s topic stamps. When running over a set of documents, switching from one document to the next is cued by a colour change and a vertical scroll.



5.2 RSVP

RSVP presents the user with a sequence of topically salient phrases as per the capsule overview analysis. In its basic mode with no user interaction, RSVP cycles through salient relational contexts in a document, maintaining the order in which they appear in the text. Each context phrase is displayed as the prominent object on the screen; at the same time the context is overlaid onto topic expansions (displayed as translucent text). This facilitates further interpretations of the context strings by the user: expansions relate phrasal contractions in the string displayed to their full canonical forms in the text, make clear antecedents of dangling anaphors, and so forth. Note, for instance, the background display of the full form of the antecedent for the anaphoric “he” in the foreground.

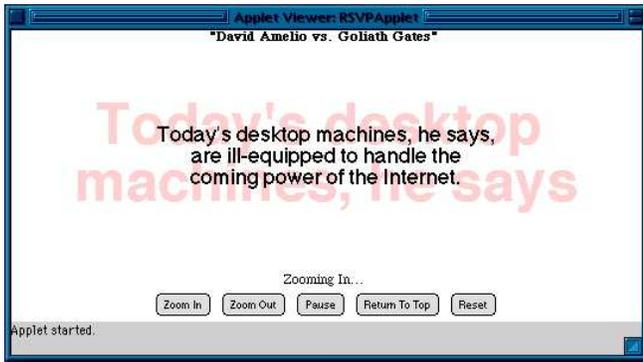


Cycling through the complete set of salient contexts, in their original sequence, offers a good indication of aboutness at a given level of depth and detail. Granularity of display is adjustable via a parameter: thus RSVP could be reduced to a topics ticker by only cycling through the document’s topic stamps, or it could be used to display sequences of sentences. Relational contexts offer just the right balance between terseness (phrases are more easily perceived and assimilated than sentences) and informativeness (phrases larger than ‘bare’ topic stamps convey richer data). The amount of time a phrase is displayed is dynamically calculated, based on studies of active reading and perception; the intent is to optimise the full document display regime so that individually, each phrase can be processed by the user, while globally, the entire set of salient contexts can be cycled through rapidly.

There are provisions for maintaining overall context, by continuously displaying the title of the current document, as well as for allowing context switching, by user selection of a document from a pop-up menu.

RSVP is designed as an entirely autonomous viewer: after all the content highlights in a document have been displayed, the next document will be loaded and the cycle repeated (just like in TOPICSTICKER, a suitable visual cue signals document change). This makes it very appropriate for situations where readers might not have much time, bandwidth, or opportunity to interact with the display, but they would want to be

peripherally aware of new documents that come into the system. On the other hand, if a particular context catches the user attention, a visual ‘zoom’ mechanism makes use of the multiple levels of analysis of the document (as defined via the containment hierarchy, see 2.2). This will reveal, on demand, progressively larger and more detailed document fragments: sentences, paragraphs and segments. For instance, further specifics concerning what “*he says*” (see the earlier illustration) is immediately available by a single click in the display area.



At any given point of time and depth of detail, the display uses a combination of visual cues to highlight the information-bearing unit which is in focus, and associate this with the larger context in which it appears in the original source. In particular, properties of type, both static and dynamic, come to convey various aspects of the document analysis: primary focus of attention is denoted by using heavy black typeface; background context, by using translucent text; semantic relatedness, by overlaying the focus and context onto the same display area; different level of attention to detail, by visually and perceptibly zooming in when more detail is requested, and by zooming out when the user retreats back into the default ‘fly-through, from bird’s eye view’ mode. Note that while such visual devices are very effective for delivering document highlights, they rely crucially on a mechanism capable of carrying out the layered analysis described in Section 3 earlier.

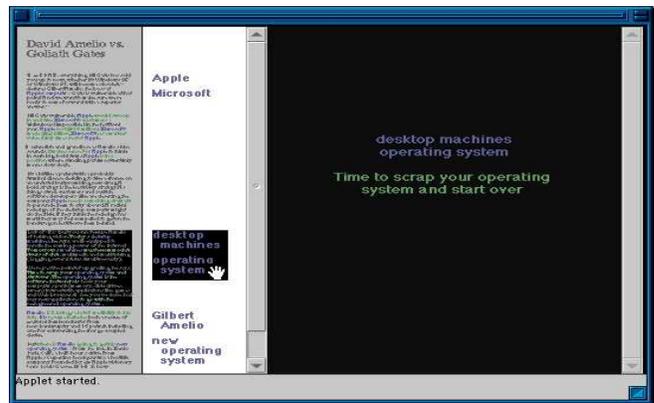
The RSVP viewer is particularly well-suited for deployment in a screen saver mode, in a background window on a desktop machine, or on a large screen projection in communal areas. In any of these situations, a topic or context might catch a reader’s peripheral attention, and then they can decide to take a further look. RSVP thus naturally extends, and fits into, the emerging ‘push’ model of information delivery.

5.3 ViewTool

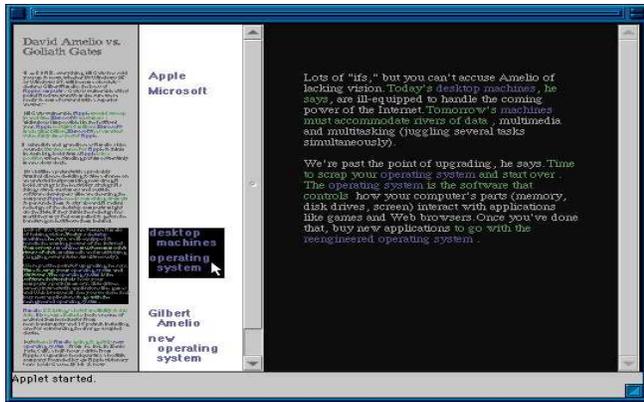
The VIEWTOOL viewer freely borrows some of the ideas of RSVP. However, the emphasis here is to present a fuller overview of the salient topic stamps in a document, contextualising this to a document ‘thumbnail’, indicative of the distribution of these highly salient objects in the text. At the same time, a separate ‘details’ area constantly displays additional information contextualising the current focus of user’s attention. The details area is used both for dynamic display of richer contexts, as in the RSVP viewer, and for providing access to the full text, or topically coherent segments from

it, on demand. Thus, the aim of this viewer is to develop a more elaborate notion of context, while maintaining permanent focus on the salient highlights (topic stamps) in the document. VIEWTOOL further seeks to offer more interactivity to the user, in ways which make the targeted exploration of portions of the document natural and transparent.

VIEWTOOL places the capsule overview of a document within the context of the document itself. This is maintained by synchronized display of discourse segments, topic stamps, and relational contexts in three panels. The whole document is displayed in the left panel; this is deliberately unreadable, and is intended to function as a document thumbnail serving as a contextual referent for the topics presented in the central panel. With the use of an appropriate colour coding scheme, it also serves as an indicator of the distribution of topically prominent phrases in the document. The central panel lists the highly salient topic stamps. Contextualisation for these is achieved by aligning the topic stamps for a given discourse segment with the textual span of that segment in the thumbnail. This offers an immediate overview of, for instance, what is being discussed in the beginning of the document, or in the end, or which topics keep recurring throughout, and so forth.



The central panel is sensitive to the user’s focus of attention: as the mouse rolls over a topic stamp, the discourse segment from which this topic has been extracted is highlighted in the left panel. The highlighting also indicates the segmentation of the source documents into topically different, and distinct, text sections. This design makes it easy to do rapid selection of areas of interest in the document, as it is mediated by the topic stamps per segment display. Again, the granularity of analysis and the layered contextual information in the capsule overview make it easy to offer immediate and more detailed information about any given set of topic stamps: simultaneously with highlighting the appropriate discourse segment in the left panel, relational contexts for the same set of topic stamps are displayed cyclically, in RSVP-like fashion, in the right panel. This ensures that topic stamps are always related with contextual cue phrases. Thus an additional level of detail is made available to the user, with very little ‘prompting’ on their part; if it is still the case that the full text of the segment would be required, clicking on its ‘proxy’ topic stamps (in the middle panel) would display this in the right panel. The larger area available there, as well as an automatic readjustment of the size of type, ensures that the text is readable.



As a natural extension of the same metaphor, clicking on the document proxy in the left panel brings up the full document text in the right panel. The full text always uses colour markup to indicate, in yet another way, topically salient phrases and their relational contexts.

By always maintaining the larger document context for any grain of information of relevance and interest in the documents, VIEWTOOL is an ideal skimming tool, because it provides additional information that may be important in deciding whether looking more closely at the document would be required. For example, users can get a sense of the size of the document, whether it contains any pictures, and other visual cue features. They can see the density of topics and the relevant ordering of the topics in relation to the different sections of the document. The tool offers the ability to see arbitrarily detailed contextual information relevant to a topic, while leveraging that same containment hierarchy of layered information units to prevent overload.

6 Conclusion

The viewers are fully implemented in Java, and are deployed in a variety of ways within a suite of intranet tools for collaboration and communication within communities. To demonstrate the capabilities of the system described here, an on-line newspaper has been configured as the primary news source within a learning community. By means of a variety of Web spiders and document filters, external news stories are collected and 'published' in the newspaper. RSVP is used for primary delivery of the external news, on a dedicated page, projected on large display in a shared common area. TOPICSTICKER offers a brief overview of the latest news on the front page of the newspaper. VIEWTOOL is available as an alternative browser, for more pro-active and focused access to the document—particularly in situations where the newspaper is being viewed on personal workstations. VIEWTOOL is also used for browsing of personal information feeds, sent to a document analysis server engine via a simple e-mail protocol. In general, any configuration of viewers can be deployed for personalised 'windows' into continuous news feeds, combining a variety of screen delivery modes.

The notions of temporal typography and dynamic delivery of content mediated via content highlights offer an appealing synergy of form and content, which not only alleviates inevitable (given the current state-of-the-art of text processing

technology) shortcomings of summarisation technologies today, but also suggests that additional utility, and user satisfaction, can be derived from imperfect analysis technologies—if usability and interface issues are addressed from a novel perspective.

Acknowledgements

We would like to thank Stephanie Houde for her critiques and comments during the design phase of these viewers. We also thank two anonymous reviewers who offered detailed and helpful comments on an earlier draft.

References

- [1] Advanced Technologies Group, Apple Computer, Cupertino, CA. APPLE INFORMATION ACCESS TOOLKIT: *Developer Notes and API's*, 1997.
- [2] Association for Computational Linguistics. *Proceedings of a Workshop on Intelligent, Scalable Text Summarization*, Madrid, Spain, 1997.
- [3] R. Barzilay and M. Elhadad. Using lexical chains for text summarization. In *Proceedings of ACL'97 Workshop on Intelligent, Scalable Text Summarisation*, Madrid, Spain, 1997.
- [4] B. Boguraev and C. Kennedy. Saliency-based content characterisation of text documents. In *Proceedings of ACL'97 Workshop on Intelligent, Scalable Text Summarisation*, Madrid, Spain, 1997.
- [5] U. Hahn and M. Strube. Centered segmentation: scaling up the centering model to global discourse structure. In *Proceedings of ACL-EACL'97, 35th Annual Meeting of the Association for Computational Linguistics and 8th Conference of the European Chapter of the Association for Computational Linguistics*, Madrid, Spain, 1997.
- [6] M. Hearst. Multi-paragraph segmentation of expository text. In *32nd Annual Meeting of the Association for Computational Linguistics*, Las Cruces, New Mexico, 1994.
- [7] M. A. Hearst. Tilebars: Visualization of term distribution information in full text information access. In *ACM SIGCHI Conference on Human Factors in Computing Systems*, Denver, CO, 1995.
- [8] E. Hovy and C. Y. Lin. Automated text summarization in SUMMARIST. In *Proceedings of ACL'97 Workshop on Intelligent, Scalable Text Summarisation*, Madrid, Spain, 1997.
- [9] N. Hutheesing. Gilbert Amelio's grand scheme to rescue Apple. *Forbes Magazine*, December 16, 1996.
- [10] F. Karlsson, A. Voutilainen, J. Heikkilä, and A. Antilla. *Constraint grammar: A language-independent system for parsing free text*. Mouton de Gruyter, Berlin/New York, 1995.
- [11] C. Kennedy and B. Boguraev. Anaphora for everyone: Pronominal anaphora resolution without a parser. In *Proceedings of COLING-96 (16th International Conference on Computational Linguistics)*, Copenhagen, DK, 1996.

- [12] C. Kennedy and B. Boguraev. Anaphora in a wider context: Tracking discourse referents. In W. Wahlster, editor, *Proceedings of ECAI-96 (12th European Conference on Artificial Intelligence)*, Budapest, Hungary, 1996. John Wiley and Sons, Ltd, London/New York.
- [13] K. Mahesh. Hypertext summary extraction for fast document browsing. In *Proceedings of AAAI Spring Symposium on Natural Language Processing for the World Wide Web*, pages 95–104, Stanford, CA, 1997.
- [14] D. Marcu. From discourse structures to text summaries. In *Proceedings of ACL'97 Workshop on Intelligent, Scalable Text Summarisation*, Madrid, Spain, 1997.
- [15] R. Rao, J. O. Pedersen, M. A. Hearst, J. D. Macinlay, S. K. Card, L. Masinter, P.-K. Halvorsen, and G. G. Robertson. Rich interaction in the digital library. *Communication of the ACM*, 38(4):29–39, 1995.
- [16] K. Sparck Jones. Summarising: Where are we now? Where should we go? In *Keynote address to ACL'97 Workshop on Intelligent, Scalable Text Summarisation*, Madrid, Spain, 1997.
- [17] Y. Y. Wong. Temporal typography. Characterization of time-varying typographic forms. Master's thesis, MIT Media Lab, 1995.