Generating Coherent Extracts of Single Documents Using Latent Semantic Analysis

by

Tristan Miller

A thesis submitted in conformity with the requirements
for the degree of Master of Science
Graduate Department of Computer Science
University of Toronto

Abstract

Generating Coherent Extracts of Single Documents
Using Latent Semantic Analysis

Tristan Miller
Master of Science
Graduate Department of Computer Science
University of Toronto
2003

A major problem with automatically-produced summaries in general, and extracts in particular, is that the output text often lacks textual coherence. Our goal is to improve the textual coherence of automatically produced extracts. We developed and implemented an algorithm which builds an initial extract composed solely of topic sentences, and then recursively fills in the lacunae by providing linking material from the original text between semantically dissimilar sentences. Our summarizer differs in architecture from most others in that it measures semantic similarity with latent semantic analysis (LSA), a factor analysis technique based on the vector-space model of information retrieval. We believed that the deep semantic relations discovered by LSA would assist in the identification and correction of abrupt topic shifts in the summaries. However, our experiments did not show a statistically significant difference in the coherence of summaries produced by our system as compared with a non-LSA version.

Acknowledgments

I would like to thank my supervisor, Graeme Hirst, for his advice and support. Thanks are also due to my second reader, Gerald Penn, for his thoughtful comments and suggestions. My gratitude is extended to the University of Toronto Statistical Consulting Service for having suggested appropriate analyses for my experimental data. Finally, acknowledgment is made to the Natural Sciences and Engineering Research Council, which has provided financial support for this research.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

It has been rightly said that we live in an age of information overload. According to Pratkanis and Aronson [2001, p. 3], the United States market alone boasts 1509 daily newspapers, 7047 weekly newspapers, and over 17 000 magazines and newsletters; over 50 000 new books are published each year. The glut of online information is no less astounding—Lawrence [2001] estimates that over a million scientific articles are currently available on the World Wide Web, and at the time of this writing, the Google search engine indexes some three billion Web pages and 700 million Usenet postings. The increasing availability and quantity of information has created an important opportunity for broad application of information processing tools.

Professionals in fields from journalism to health care to scientific research deal with the time-consuming task of information management every day. Distilling the ideas contained in large documents, or in groups of related documents, simplifies the task of searching, classifying, and organizing information. Using a brief summary of a large text, a researcher can often tell at a glance whether the document is relevant to his or her work, and can easily identify the most salient points addressed by the text. Our goal is to develop a general-purpose summarization system which not only extracts the most relevant ideas from a single source document, but also presents them in as clear and coherent a manner as possible.

In this chapter, we provide a brief overview of the field of automatic summarization and the technology behind the system we have developed. Chapter 2 describes in detail the architecture and implementation details of our summarization system. In Chapter 3, we describe our methodolgy for evaluating the system's performance, and present and discuss the results obtained from our experiment. Finally, Chapter 4 offers further insight into the meaning of the results and suggests some avenues for future research.

1.1 Basic notions

1.1.1 Summarization

Summarization is the process by which the most important concepts in a document are identified and then presented in a condensed, human-readable form. Human-produced summaries usually contain text which is not found in the source document; such summaries are called *abstracts*. Due to the difficulty of automating natural language understanding and generation, however, automatic summarization is usually reduced to the task of *extraction*, where the summary consists of text taken verbatim from the input document. In this paper we concern ourselves with the production of extracts only, so except where noted, we henceforth use the terms *summary* and *extract* interchangeably.

The size of a summary relative to its source document is known as its rate of *compression* or *condensation*, and is usually expressed as a percentage of the input size (measured in characters, words, sentences, or some other suitable unit). Thus a summary with 10% compression is comparatively small and is said to be "highly" compressed. "Lightly" compressed summaries (say, 50–90%) are known as *digests*. The rate of compression is determined by the user's needs, which may be influenced by the format and topic of the data to be summarized. In this thesis we focus on the task of generating general-purpose summaries of single documents in any topic domain.

The ideal summary is one which captures all the salient information in the source document and presents it in a clear, coherent manner. Unfortunately, defining these criteria precisely is not a simple task. Mani [2001, p. 11] defines *salience* or *relevance* as "the weight attached to information in a document, reflecting both the document content as well as the relevance of the document information to the application". A summary that maximizes inclusion of relevant information from the source is said to be *comprehensive*.

Coherence is the way the parts of the text gather together to form an integrated whole; a coherent text is one which is well-organized and has no confusing gaps in the reasoning or progression of ideas. Coherence is often difficult to distinguish from *cohesion*, which refers to the "connectedness" of text and is determined by relationships (often grammatical) between words and referring constructs, such as repetition, anaphora, and conjunctions. Take the following example:

> John's car won't run. Its engine is shot.

This sentence pair exhibits both coherence and cohesion. The most obvious cohesive tie between the two sentences is the use of the pronoun "its" in the second sentence to refer to "John's car" in the first. There is also a coherence relation between the two sentences, which Hobbs calls "elaboration"[1] because the second sentence provides further information to substantiate or explain the claims in the first. It is difficult to point to a single defining feature of this sentence pair which allows us to say that there is a coherence relation. Certainly part of what allows us to see that there is an elaboration involves our real-world knowledge of the strong meronymy relation between "car" and "engine" which, it could be argued, is simply an instance of lexical cohesion. However, it is certainly possible to construct elaborations where the semantic relations across sentences are not so easily classified. For example, take the following elaboration:

[1] There is no consensus among researchers on how to name and classify coherence relations. As we shall see in §4.2.1, however, recent work may change this.

The secretary cannot work. Her typewriter is broken.

There is no clearly-identifiable lexical relationship between "secretary" and "typewriter", let alone a three-way relationship among "secretary", "work", and "typewriter". Yet on an intuitive level, we know that these relationships exist, and we must use them to infer the coherence relation.

Another important criterion for summaries is the degree of *redundancy*. While a certain amount of repetition is necessary for fluent communication [Irwin, 1980], rote reiteration of the same or similar sentences is rarely useful. Particularly in highly-compressed summaries, redundant sentences take up precious space that should have been filled with novel information.

1.1.2 Information retrieval

Vector-space model

Much of the current research in automatic summarization, including ours, draws from related work in the field of information retrieval (IR). Of particular relevance is many systems' use of the vector-space model [Salton and McGill, 1983] to measure, or at least approximate the measurement of, semantic content. In the original IR model, a set of documents is conceptualized as a two-dimensional co-occurrence matrix, where the columns represent the documents and the rows represent the unique terms (usually words or short phrases) occurring in the documents. Sometimes every term appearing in the source document will be represented by a row, though it is more common to exclude a *stop list* of prepositions, function words, and other lexemes with negligible semantic content. The value in a particular cell may be a simple binary 1 or 0 (indicating the presence or absence of the term in the document) or a natural number indicating the frequency with which the term occurs in the document. Typically, each cell value is adjusted with an information-theoretic transformation. Such transformations, widely used in IR (*e.g.,* Spärck Jones, 1972), weight terms so that they more properly reflect their importance within the document. For example, one popular measure known as TF–IDF (term frequency–inverse document frequency) uses the following formula:

$$w_{ij} = \mathit{tf}_{ij} \log_2 \frac{N}{n_i}.$$

Here w_{ij} is the weight of term i in document j, tf_{ij} is the frequency of term i in document j, N is the total number of documents, and n_i is the number of documents in which i occurs. After the weighting, pairs of documents can be compared by their column vectors, using some mathematical measure of vector similarity. Perhaps the most popular measure is the cosine coefficient,

$$\cos(A, B) = \frac{\sum_i A_i B_i}{|A| \cdot |B|}.$$

Some automatic summarization systems use the vector-space model to compare the semantic similarity of discourse units within a single document. In this case, the "documents" of the term–document co-occurrence matrix are actually sentences or paragraphs.

Latent semantic analysis

Latent semantic analysis, or LSA [Deerwester *et al.*, 1990; Landauer *et al.*, 1998], is a technique originally developed for solving the problems of synonymy and polysemy in information retrieval. Its basic assumption is that every document has an underlying semantic structure, and that this structure can be captured and quantified in a matrix. LSA is unusual among natural language processing techniques in that it makes no use of human-constructed parsers, taggers, dictionaries, semantic networks, or other tools. The input is simply a collection of documents separated into words or meaningful terms.

LSA is based on the vector-space model discussed previously, but it extends the model in a very important way. Specifically, it exploits *singular value decomposition*, a well-known theorem in linear algebra which asserts that any real-valued rectangular matrix,

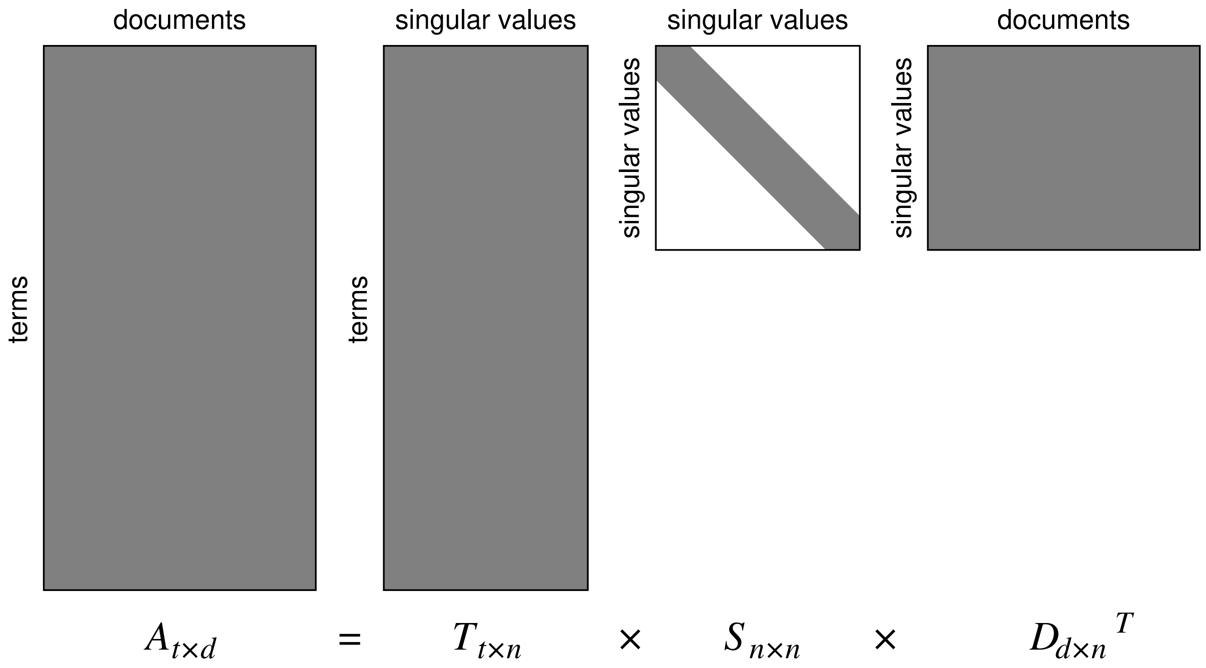

Figure 1.1: Singular value decomposition of a matrix

such as a term–document co-occurrence matrix of the form previously described, can be represented as the product of three smaller matrices of a particular form. The first of these matrices has the same number of rows as the original matrix, but has fewer columns. These n columns correspond to new, specially derived factors such that there is no correlation between any pair of them—in mathematical terms, they are linearly independent. The third matrix has the same number of columns as the original, but has only n rows, also linearly independent. In the middle is a diagonal $n \times n$ matrix of what are known as *singular values*; without loss of generality, these values are monotonically non-increasing. The purpose of the singular value matrix is to scale the factors in the other two matrices such that when the three are multiplied, the original matrix is perfectly reconstructed. Figure 1.1 illustrates the decomposition of a term–document matrix A with t distinct terms and d documents into three constituent matrices T, S, and D^T.

Things get more interesting, however, when *fewer* than the necessary number of factors are used to reconstruct the original matrix. This can be done by deleting (*i.e.,* setting to zero) one or more of the smallest values from the singular value matrix, which causes

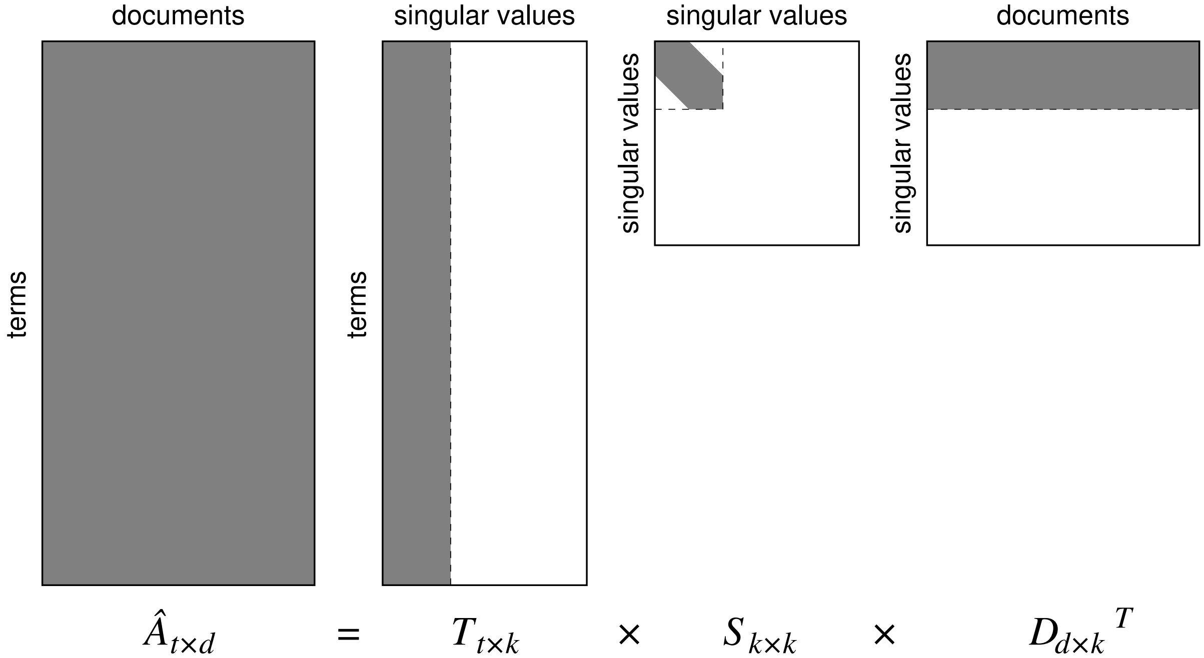

Figure 1.2: Approximate recomposition of a matrix

the same number of columns and rows from the first and third matrices, respectively, to be disregarded during multiplication. In this case, the product of the three matrices turns out to be a least-squares best fit to the original matrix. Figure 1.2 illustrates this procedure; here, the $n-k$ smallest singular values have been deleted from S, as indicated by the dashed line. This effectively causes the dimensionality of T and D^T to be reduced as well. The new product, $\hat{A}$, still has t rows and d columns, but is only approximately equal to the original matrix A.

Taken in the context of a term–document co-occurrence matrix, this means that many terms may appear with greater or lesser frequency in the reconstructed matrix than they did originally. In fact, certain terms may appear at least fractionally in documents they never appeared in at all before. The apparent result of this smearing of values is that the approximated matrix has captured the latent transitivity relations among terms, allowing for identification of semantically similar documents which share few or no common terms withal.[2] For example, assume that a collection has some documents that

[2]Likewise, terms may be compared by examining their vectors across documents. Terms may be

contain the terms *animal* and *dog*, and some documents that contain the terms *animal* and *hound*. Furthermore, assume that *hound* never occurs in a document containing *dog*, and vice versa. Even though *dog* and *hound* never co-occur, the strength of their statistical association will be reflected in the LSA matrix. Using the cosine metric, the *dog* documents will be found semantically similar to the *hound* documents, and probably significantly more so than to those documents containing *animal* alone. Kontostathis and Pottenger [2002] report that LSA is able to infer not only second-order semantic relations such as this one, but also third-, fourth-, and fifth-order relations. The usefulness of this property becomes apparent when one considers that two people will use the same word for a well-known referent less than 20% of the time [Furnas *et al.*, 1983]. For instance, the United States of America is variously referred to as *America*, *the US*, *the USA*, *the United States*, and *the States*.

There is one well-known problem with LSA, which is that determining the number of dimensions by which to reduce the scaling matrix is somewhat of a black art. Too little a reduction reconstructs the original matrix too faithfully to capture any latent semantic information; too large a cut renders the matrix too noisy to be useful. The optimal dimensionality must be determined empirically. Once a suitable degree of reduction has been discovered, however, two documents can be compared in time linear to the number of terms.

1.2 Related work

1.2.1 Generating coherent summaries

The earliest work in automatic summarization is that of Luhn [1958], wherein extracts were created by selecting sentences containing *content words*. These words were found by compiling a frequency list of all words appearing in the document and then removing the

judged semantically similar even though they never occur in the same text together.

words beyond high- and low-frequency cutoffs, as well as those contained in a stop list. The work of Edmundson [1969] significantly expanded on this approach by considering the document's structure as well; words were weighted on the basis of their position within the sentence and whether they occurred in the document title or section headings.

A major problem with these early approaches, and indeed with almost every other extract-based system since developed, is that the output text often lacks fluency and organization. Sentences often leap incoherently from topic to topic, confusing the reader and hampering his ability to identify information of interest. Interest in producing coherent summaries has consequently increased in recent years, leading to a wide variety of approaches. The earliest and simplest techniques exploited the correlation between cohesion and coherence by enforcing the former to achieve the latter. For example, if an extractor selects a sentence containing an pronoun, it might automatically select the previous sentence in hopes of providing the referent as well, preserving not only the cohesive tie but also the flow of ideas. The problems with this naïve approach become apparent when one considers that not all pronouns are anaphora, that those that are do not always have their referent in the preceding sentence, and that the preceding sentence may contain little other information of value.

In the following sections we present a brief overview of some of the recent work in generating coherent summaries, plus some older projects which we consider to be similar to our own approach. Unfortunately, very few of the papers describing the systems we review contain evaluation measurements for textual coherence, so we must take the authors' words that these approaches have the intended effect.

IR-based techniques

Carbonell and Goldstein [1998] draw from work in IR for their maximum marginal relevance (MMR) summarizer, which produces summaries tailored to a particular topic or point of view as embodied in a query string. Using the cosine similarity coefficient, they rank all sentences in the source document on the basis of their similarity to the query string and to each other. Highly-ranked sentences have maximal similarity to the query string and minimal similarity to all other sentences in the summary. More formally,

$$MMR = \underset{D_i \in R \setminus S}{\arg\max}\, \lambda \left(\cos(D_i, Q) - (1 - \lambda) \max_{D_j \in S} \cos(D_i, D_j) \right),$$

where Q is the query string, R is the set of sentences in the source document, S is the subset of sentences in R already selected for inclusion in the summary, cos() is the cosine similarity metric, and the weighting $\lambda \in [0, 1]$ determines whether the sentence selection should be biased more in favour of relevance to the query string ($\lambda < 0.5$) or to maximal sentence diversity ($\lambda > 0.5$). For higher values of λ, returning the n top-ranked sentences in their original order of appearance yields an extract with the broadest possible topic coverage (relative to the query string) with a minimum of redundancy. Though they do not claim it to be a goal of their summarizer, good textual coherence may be a consequence of their ranking criteria, since after the maximally diverse sentences have been selected, overall intersentential similarity increases as more sentences are added.

The method of Salton *et al.* [1997] is also IR-based, but does not depend on a query string. They first compare all paragraphs in the source document to each other using the cosine similarity coefficient. Topic boundaries are identified by finding paragraphs which compare well with successive paragraphs but poorly with preceding ones. An extract is then constructed in the following manner:

1. Using the compression rate, determine the number of paragraphs to be taken from each segment. The number should be proportional to the segment's length.

2. Begin selecting paragraphs in order of their similarity to all other paragraphs in the document. (Such paragraphs are said to be *bushy*, because their corresponding nodes in a connectivity map have high degree.) Once a paragraph is selected for use, no paragraphs occurring earlier in the document may be selected in this step.

3. For each segment, start at an important paragraph in the segment (*e.g.,* the first paragraph, or a highly bushy node) and construct a path of paragraphs $p_1, p_2, \ldots$ such that p_{i+1} is in the segment and has the highest similarity to p_i of any paragraph occurring after p_i. Add the path to the extract.

4. For each segment except the last one, supply "transition paragraphs"—*i.e.,* those exhibiting high similarity to the initial paragraphs of the succeeding segment.[3]

5. Present the selected paragraphs in their order of appearance in the original document.

The second and third steps are meant to ensure comprehensiveness, and the fourth and fifth steps, coherence.

Unfortunately, due to the use of paragraphs as the discourse unit, this method is prone to using up available space very quickly. (The method does not work well with just sentences due to the paucity of common terms to be matched by the cosine metric.) By predetermining the amount of text to extract from each topic segment, the method fails to account for redundancy, or lack thereof, in the source document. Highly redundant segments will receive an unfairly large proportion of summary space, possibly robbing abrupt topic shifts of the transition material they need. Furthermore, the algorithm may add transition paragraphs where none are really necessary.

Lexical chaining techniques

Exploiting the close relationship between cohesion and coherence, the University of Lethbridge summarizer [Brunn *et al.*, 2002] attempts to generate fluent summaries of single or multiple documents using lexical chaining [Morris and Hirst, 1991]. A lexical chain is, in essence, a chain of words in a text such that each word in the chain bears some kind

[3] It is not clear from the article whether this step is mandatory (*i.e.,* space is reserved for at least one transition paragraph), or is performed only when there is still space to be filled in the segment after completion of the previous two steps.

of cohesive relationship (hyponymy, meronymy, *etc.*) to a word that is already in the chain. In the Lethbridge algorithm, the source text is first segmented into discrete topics using the C99 algorithm [Choi, 2000], and lexical chains are computed for each segment. Each segment is then assigned a score based on the ratio of chain members occurring in the segment to the number of segments in which those chain members occur. Sentences which contain large numbers of words belonging to their segment's lexical chains are then extracted from the highest-ranking topic segments. Because lexical chains are essentially strings of related words characterizing a particular topic, this approach may fail to bridge coherence gaps *between* topics. The magnitude of this problem depends in part on the granularity of the topic segmenter.

Karamuftuoglu [2002] has experimented with a related technique. He defines a *lexical link* between two sentences as a word stem that occurs in both, and a *lexical bond* to be two or more lexical links. An SVM-based machine learning system [Vapnik, 1995] is used to select sentences for the extract; the feature set includes the number of lexical links, the number of forward and backward lexical bonds, and various other surface linguistic features. The resulting summaries are remarkably coherent, but at too great a price: in his trials, 58% of the summaries were formed by sentences that sequentially follow each other in the original text. Entire sections of the original document, often containing important topics, are completely overlooked. Furthermore, particularly for short summaries, sentences occurring early in the source document are grossly overrepresented.

Karamuftuoglu [2002] also discusses an alternative approach wherein extracts are produced by selecting the first sentence in the document which has a forward lexical bond, and then following the chain of forward lexical bonds from one sentence to the next. The problem is that a sentence may have lexical bonds with more than one subsequent sentence; the branching factor was found to be unmanageable even after the introduction of additional selection constraints.

Discourse structure techniques

Mani [2001] lists several established analyses of argument structure which could conceivably be used to address textual coherence issues in NLP. These analyses include rhetorical structure theory (RST) [Mann and Thompson, 1987, 1988], discourse grammar [Longacre, 1979], macrostructures [van Dijk, 1988], and coherence relations [Hobbs, 1985]. With the exception of RST, however, little work in automatic summarization has been done with these analyses, in large part because they were never designed as computational models. Much of the work in automatic summarization that does incorporate RST (*e.g.,* Marcu, 1997, 1999; Chan *et al.*, 2000) sees textual coherence as a means rather than an end; that is, coherence relations in the source text are identified and classified only in order that the most salient concepts therein may be extracted. The summaries themselves are not guaranteed to read smoothly. However, because RST-based systems discover a great deal about the discourse structure of the source text, it is conceivable that, paired with natural language generation techniques, future research could put this information to use in generating coherent abstracts.

Other techniques

Recent work on the RIPTIDES system [White and Cardie, 2002; White *et al.*, 2002] is similar to ours in that summary coherence has been made a top priority. The authors view sentence extraction from multiple documents as a randomized local search procedure [Selman and Kautz, 1994] where the selection of adjacent sentences is rewarded and the inclusion of redundant material is penalized. Though the description of their algorithm suggests that any textual coherence it produces is merely a byproduct of cohesion-preserving techniques, we discuss it here because the literature includes comparatively extensive evaluations of coherence.

The basic algorithm for RIPTIDES is as follows:

1. Perform surface-oriented clustering to group together sentences which address a

common topic.

2. Score the sentences in the original documents by considering some weighted combination of surface features, position within the document, and semantic similarity to other sentences. (Semantic similarity is measured with Columbia University's SimFinder tool [Hatzivassiloglou *et al.*, 2001].)

3. Create an initial summary by selecting the highest-scoring sentences.

4. Repeat the following n times:

   (a) Score the summary as follows:

      i. The base score is the sum of the scores of the sentences in the summary.

      ii. Penalize for inclusion of multiple sentences from the same topic cluster.

      iii. Penalize for inclusion of sentences whose similarity exceeds a certain threshold.

      iv. Reward inclusion of sentence pairs which are adjacent in the original documents, more so if

         - the second sentence begins with a pronoun, or
         - the second sentence begins with a discourse marker (*e.g., however*).

   (b) Randomly select and perform one of the following steps until a greedy step fails to improve the summary score:

      Random step. Randomly select a sentence from the source documents and add it to the summary.

      Greedy step. Add one sentence to the summary, and remove zero or more sentences, such that the summary size remains under the limit specified by the compression ratio, and the new combination of sentences represents the best swap according to the summary-scoring scheme.

(c) Create a new summary composed of sentences chosen at random from the source documents.

5. At this point, the algorithm will have generated n summaries, $n - 1$ of which were produced from random starting points. Select the highest-scoring of the n summaries as the final version.

It has been found that $n = 10$ produces acceptable summaries in under a minute of computation time.

For the evaluation phase, the authors had two human judges rank summaries produced by six systems with respect to content and intelligibility.[4] The systems included the standard RIPTIDES system and three simpler versions of it, including a simple marginal relevance system inspired by Carbonell and Goldstein [1998]. There were also two naïve baselines: the initial sentences of the latest article in the document set, and the paragraph-initial sentences of the latest article in the document set (in both cases only up to the summary length limit). The first baseline consistently ranked the highest for intelligibility, while the second baseline and the marginal relevance system were almost always the two lowest-ranked. It is not clear whether the remaining systems are statistically distinguishable.[5]

1.2.2 Related semantic analysis approaches

The Sumatra summarization system [Lie, 1998] employs a semantic analysis component which, like LSA, attempts to identify latent semantic relations among the terms of the source document. Unlike LSA, however, the technique is not based on the vector-space model. Instead, Sumatra uses a sort of unlabelled dependency tree as its basic semantic

[4] In a personal communication from the authors, it was explained that the content rank was meant to assess relevance, and intelligibility "coherence, cohesion, and also repetitiveness".

[5] Actually, it is not clear whether any of the results obtained are statistically valid. The two judges were the authors themselves, and the evaluation, while conducted blind, always presented the summaries generated by the various systems in the same order.

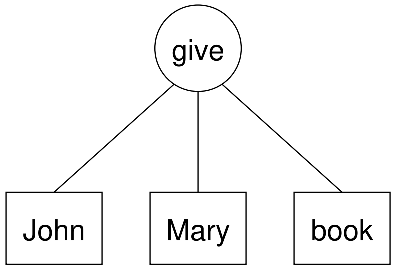

Figure 1.3: Dependency tree in Sumatra

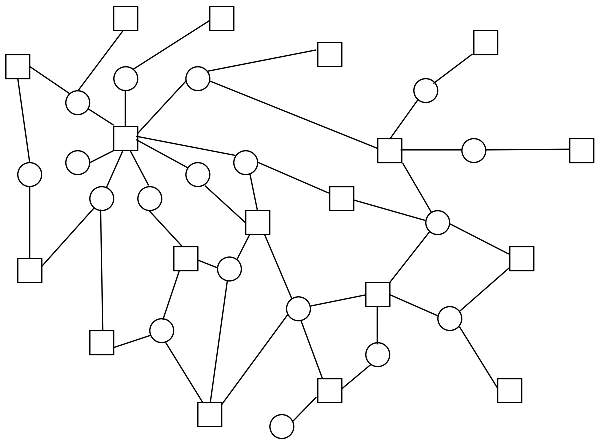

Figure 1.4: Semantic network in Sumatra

unit, and conceptualizes a document's semantic structure as the graph created by the union of all such trees. Figure 1.3 illustrates how the sentence "John gives Mary a book" would be converted to a dependency tree, and Figure 1.4 depicts a graph resulting from the union of many such trees, where each object and relation type is represented by a unique node. Important concepts in the source document correspond to areas of high connectivity in the semantic network. Exactly how these subgraphs are converted back to sentences is not made clear by the available literature. Lie [1998] claims that the system uses natural language generation and aggregation techniques to produce text directly from the graphs, but judging from the summaries produced by the system we obtained (see §3.1.2), Sumatra is a simple sentence extractor.

1.3 Research statement

We propose a new, iterative method for automatic text summarization which attempts to preserve both the comprehensiveness and the coherence of the source document. Textual coherence is an important aspect of summary quality, but in highly-compressed summaries, it often comes at the expense of topic coverage. Partly for this reason, it is often overlooked by summarization researchers. However, there are some cases where preserving textual coherence may be given higher priority. For example, digests, because they retain so much of the source document, are unlikely to suffer from a lack of topic coverage. In digesting it is better to excise from the source document that which is irrelevant or repetitious, and ensure that the resulting gaps do not diminish readability. Textual coherence is also important for certain document types, as we discuss in §4.2.2.

Our system fits within the general category of IR-based systems, but attempts to circumvent some of the limitations and disadvantages of the systems previously described. Unlike in Salton *et al.* [1997], textual redundancy within topic segments will be minimized, as the contribution of each segment to the extract is determined dynamically. Moreover, relevant text will be extracted with significantly more precision through the use of LSA, which can accurately compare much smaller discourse units [Deerwester *et al.*, 1990]. Our summarizer will not require the use of a query string as does MMR [Carbonell and Goldstein, 1998], though we do not feel that it would be difficult to adapt our technique to produce query-focused summaries.

We feel that Karamuftuoglu [2002] was on the right track with his lexical-bond approach to tracking the flow of topics. However, we feel that our use of LSA rather than the simple presence of common word stems will allow for identification of much deeper and less apparent semantic relations between sentences. Karamuftuoglu [2002] tried (unsuccessfully) a top-down approach to navigating the source document's graph of lexical bonds, starting at a sentence early in the document and trying to find a path of sentences which characterizes a coherent summary. We feel that a bottom-up approach is much more manageable; we shall identify a number of nodes—topically-relevant sentences occurring anywhere in the document—and then attempt to find intermediary nodes which semantically link them together. High branching factors will no longer be an issue.

Finally, though we have acknowledged that RST-based methods may be applied to the problem of summary coherence, such methods are tied to a particular language, requiring resources such as a list of discourse cue words and a marked-up training corpus. Our technique has the advantage of not requiring any corpora, or any language-specific NLP tools besides simple word- and sentence-boundary detection routines.

Chapter 2

Summarizer

2.1 Algorithm

Our summarizer has the pipeline architecture shown in Figure 2.1. The input is a plain text document, which is converted into a list of tokenized sentences.[1] A tokenizer (*e.g.,* Grefenstette and Tapanainen, 1994; Baldwin *et al.*, 1995) and sentence-boundary disambiguation algorithm (*e.g.,* Palmer and Hearst, 1994; Reynar and Ratnaparkhi, 1997) may be used for these first steps.

The list of m sentences (indexed from 1 to m) is then segmented into linearly discrete

[1]Paragraphs could also serve as the elementary unit of discourse; this might be appropriate for summarizing very long documents.

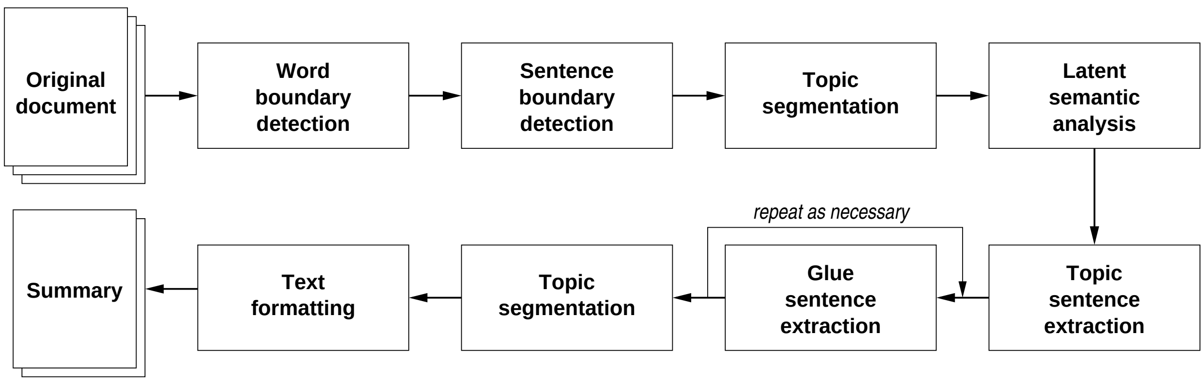

Figure 2.1: Summarizer architecture

topics. This can be done manually if the original document is structured (*e.g.,* a book with chapters, or an article with sections), or a linear text segmentation algorithm, such as C99 [Choi, 2000] or TextTiling [Hearst, 1997], can be used. The output of this step is a list of sentence indices $\langle t_1, \ldots, t_{n+1} \rangle$, where, for the ith of the n topics, t_i is the index of the first sentence of the topic segment and $t_{i+1} - 1$ is the index of the last sentence of the topic segment. We stipulate that there are no sentences which do not belong to a topic segment, so for all t_i, we have $t_i < t_{i+1}$, and

$$t_i = \begin{cases} 1 & \text{if } i = 1; \\ m+1 & \text{if } i = n+1; \\ \text{the index of the first sentence of the } i\text{th topic} & \text{otherwise.} \end{cases}$$

As mentioned previously, we use LSA to measure semantic similarity, so before we can begin constructing the extract, we need to construct a reduced-dimensionality term–sentence co-occurrence matrix. Once this is done, a preliminary extract is produced by choosing a representative "topic sentence" from each segment—that is, that sentence which has the highest semantic similarity to all other sentences in its topic segment. These topic sentences correspond to a list of sentence indices $\langle r_1, \ldots, r_n \rangle$ such that

$$r_i = \operatorname*{arg\,max}_{t_i \le j < t_{i+1}} \sum_{k=t_i}^{t_{i+1}-1} \operatorname{sim}(j,k),$$

where $\operatorname{sim}(x,y) \in [-1,1]$ is the LSA cosine similarity score for the sentences with indices x and y. In order to preserve important information which may be found at the beginning of the document, and also to account for the possibility that the document contains only one topic segment, we always consider the first sentence of the document to be a topic sentence—*i.e.,* $r_0 = 1$—and include it in our initial extract.[2] Let us refer to this initial

[2] In practice, it may be the case that $r_1 = 1$, in which case inclusion of r_0 is not necessary. For the purposes of illustration, however, we assume, without loss of generality, that $r_1 \neq 1$.

extract as $E_0 = \langle e_{0,1}, \ldots, e_{0,n+1} \rangle$ where $e_{0,i} = r_{i-1}$.

As we might imagine, this basic extract will have very poor coherence, since every sentence addresses a completely different topic. However, we can improve its coherence by selecting from the set $\langle 1, \ldots, m \rangle \setminus E_0$ a number of indices for "glue" sentences between adjacent pairs of sentences represented in E_0. We consider an appropriate glue sentence between two others to be one which occurs between them in the source document, and which is semantically similar to both. Thus we look for sentence indices $G_1 = \langle g_{1,1}, \ldots, g_{1,n} \rangle$ such that

$$g_{1,i} = \underset{e_{0,i} < j < e_{0,i+1}}{\arg\max} \; f\left(\mathrm{sim}'\left(j, e_{0,i}\right), \mathrm{sim}'\left(j, e_{0,i+1}\right)\right),$$

where

$$f\left(x, y\right) = xy \cdot \left(1 - |x - y|\right)$$

and

$$\mathrm{sim}'\left(x, y\right) = \begin{cases} 0 & \text{if } \mathrm{sim}\left(x, y\right) > \alpha \text{ or } \mathrm{sim}\left(x, y\right) < 0; \\ \mathrm{sim}\left(x, y\right) & \text{otherwise.} \end{cases}$$

for $\alpha \in [0, 1]$. The purpose of $f()$ is to reward glue sentences which are similar to their boundary sentences, but to penalize if the similarity is too biased in favour of only one of the boundaries. (See Table 2.1.) The revised similarity measure $\mathrm{sim}'()$ ensures that we do not select a glue sentence which is nearly equivalent to any one boundary—such a sentence is redundant. (Of course, useful values of α will be 1 or close thereto.)

Once we have G_1, we can construct a revised extract $E_1 = \langle e_{1,1}, \ldots, e_{1,2n+1} \rangle = \langle E_0 \cup G_1 \rangle$.[3] More generally, however, we can repeat the gluing process recursively, using E_i to generate G_{i+1}, and hence E_{i+1}. The question that arises, then, is when to stop. Clearly there will come a point at which some $e_{i,j} = e_{i,j+1} - 1$, thus precluding the

[3]For notational convenience, we take it as understood that the sentence indices in the extracts E_i are sorted in ascending order—that is, $e_{i,j} < e_{i,j+1}$ for $1 \leq j < |E_i|$.

x / y	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.10	0.00	0.01	0.02	0.02	0.03	0.03	0.03	0.03	0.02	0.02	0.01
0.20	0.00	0.02	0.04	0.05	0.06	0.07	0.07	0.07	0.06	0.05	0.04
0.30	0.00	0.02	0.05	0.09	0.11	0.12	0.13	0.13	0.12	0.11	0.09
0.40	0.00	0.03	0.06	0.11	0.16	0.18	0.19	0.20	0.19	0.18	0.16
0.50	0.00	0.03	0.07	0.12	0.18	0.25	0.27	0.28	0.28	0.27	0.25
0.60	0.00	0.03	0.07	0.13	0.19	0.27	0.36	0.38	0.38	0.38	0.36
0.70	0.00	0.03	0.07	0.13	0.20	0.28	0.38	0.49	0.50	0.50	0.49
0.80	0.00	0.02	0.06	0.12	0.19	0.28	0.38	0.50	0.64	0.65	0.64
0.90	0.00	0.02	0.05	0.11	0.18	0.27	0.38	0.50	0.65	0.81	0.81
1.00	0.00	0.01	0.04	0.09	0.16	0.25	0.36	0.49	0.64	0.81	1.00

Table 2.1: $f(x, y)$ for $x, y \in [0..1]$

possibility of finding any further glue sentences between them. We may also encounter the case where for all k between $e_{i,j}$ and $e_{i,j+1}$, $f(\mathrm{sim}'(k, e_{i,j}), \mathrm{sim}'(k, e_{i,j+1}))$ is so low that the extract's coherence would not be significantly improved by the addition of an intermediary sentence. Or, we may find that the sentences with indices $e_{i,j}$ and $e_{i,j+1}$ are themselves so similar that no glue is necessary. Finally, it is possible that the user wishes to constrain the size of the extract to a certain number of sentences, or to a fixed percentage of the original document's length. The first of these stopping conditions is straightforward to account for; the next two can be easily handled by introducing two fixed thresholds β and γ: when the similarity between adjacent sentences from E_i exceeds β, or when the value of $f()$ falls below γ, no glue sentence is suggested for the pair in question.

The case of maximum summary length is a bit trickier. If we are not concerned about undershooting the target length ℓ, then we can simply halt the algorithm once $|E_i| \geq \ell$, and then take E_{i-1} (or E_i, if $|E_i| = \ell$) as the final extract. Most real-world applications, however, demand that we maximize the extract size. Given E_{i-1} of length $\ell - p$, the optimal extract E of length ℓ is the one which glues together the p largest gaps in E_{i-1}. That is,

$$E = E_{i-1} \cup \operatorname*{arg\,max}_{G' \subset G_i : |G'| = p} \sum_{k=1}^{\ell-1} \mathrm{sim}\left(e'_k, e'_{k+1}\right),$$

Algorithm 1: glue()

input : initial extract E, maximum extract length ℓ
output : largest coherent extract of length $\leq \ell$
precondition: $|E| < \ell$
assumption : Lists are kept sorted in ascending order. Where list elements are coordinate pairs, the sorting key is the first coordinate.

```
G ← ⟨⟩;
for i ← 1 to |E| − 1 do
    s ← sim(E[i], E[i + 1]);
    if E[i] = E[i + 1] − 1 or s > β then continue;
    g ← argmax_{E[i]<j<E[i+1]} f(sim′(j, E[i]), sim′(j, E[i + 1]));
    if f(sim′(g, E[i]), sim′(g, E[i + 1])) ≥ γ then G ← G ∪ ⟨(s, g)⟩;
end
if |G| = 0 then
    return E;
else if |E| + |G| ≥ ℓ then
    return E ∪ ⟨ x | (y, x) ∈ ⋃_{i=|E|+|G|−ℓ+1}^{|G|} G[i] ⟩;
else
    return glue(E ∪ ⟨x | (y, x) ∈ G⟩ , ℓ);
end
```

where e'_k is the kth member of $E_{i-1} \cup G'$.

A version of the gluing algorithm which takes into account all four stopping conditions is shown in Algorithm 1.

Once the final set of sentences for the extract has been selected, we send the sentences, in their original order of occurrence, to the topic segmenter. The discovered topic segments are then used by a simple text formatter to partition the summary into sections or paragraphs for easy reading.

2.2 Complexity analysis

Given an initial extract of length n, the first recursion of Algorithm 1 will add at most $n - 1$ sentences to the extract, yielding a new extract of length $2n - 1$. In general, at most $2^{i-1}n$ sentences will be added on the ith recursion, bringing the extract length to $2^i n - 1$ sentences. Therefore, to achieve an extract of length $\ell > n$, the algorithm needs to recurse at least

$$\left\lceil \log_2 \frac{\ell + 1}{n} \right\rceil$$

times. The worst case occurs when $n = 2$ and the algorithm always selects a glue sentence which is adjacent to one of the boundary sentences (with indices e_1 and e_2). In this case, the algorithm must recurse $\min(\ell, e_2 - e_1)$ times, which is limited by the source document length, m.

On each recursion i of the algorithm, the main loop considers at most $m - (2^i n - 1)$ candidate glue sentences, comparing each one with two of the $2^i n - 1$ sentences already in the extract. To simplify matters, we note that $2^i n - 1$ can never exceed m, so the number of comparisons must be, at worst, proportional to m. The comparison function, sim(), runs in time proportional to the number of word types, w, in the original document (minus the stop list, if any). Thus an upper bound on the time complexity of a naïve implementation of Algorithm 1 is $O(wm^2)$.

Running time can be cut down considerably in the general case, however. Since $\text{sim}(i, j)$ remains constant, we can save time by precomputing a triangular similarity matrix of all pairs of sentences in the document, or better yet, by using memoization (*i.e.,* caching intersentential similarity values as they are computed). The algorithm could be further improved by having the loop skip over adjacent extract sentences for which no glue was found on a previous recursion. At any rate, the running time of the summarizer as a whole will likely be dominated by the singular value decomposition step of the LSA stage (at least $O(wm^2)$) and possibly too by the topic segmenter (for C99, also $O(wm^2)$).

2.3 Implementation details

The actual implementation of our summarizer used in our experiments (described in the next chapter) was developed and run in a Unix environment. The various modules of the pipeline shown in Figure 2.1 were coordinated by a script written in KornShell 93. We implemented our own regular-expression-based word- and sentence-boundary detection routines for use with English text, and used Choi's language-neutral C99 algorithm [2000] for topic segmentation. Telcordia Technologies supplied us with their LSA suite, which was invoked by our own topic- and glue-sentence extractors written in KornShell 93.

Once the system was built, we focussed our efforts on determining the optimal dimensional reduction (see §1.1.2) and similarity cutoff thresholds α, β, and γ (see §2.1). On the basis of our own informal evaluations of the generated summaries, we found that retention of 20–30% of the singular values produced reasonably good summaries. Long documents (over 70 sentences) seemed to summarize best at the lower end of this range, and short documents at the higher end. However, performance dropped off rapidly below 15% and above 30%. We found threshold values of $\alpha = 0.9$, $\beta = 1.0$, and $\gamma = 0.1$ to be appropriate for the 20–30% range; more parsimonious cutoffs tended to result in summaries greatly in deficit of the allowed length.

Chapter 3

Evaluation

3.1 Methodology

3.1.1 Introduction

In general there are two approaches to evaluating summaries: *intrinsic* evaluations, which rate the summary in and of itself, and *extrinsic* evaluations, which test the summary in relation to some other task [Spärck Jones and Galliers, 1996]. Popular intrinsic approaches include *quality evaluation*, where human graders grade the summary in isolation on the basis of relevance, grammaticality, readability, *etc.*; and *gold-standard comparison*, where the summary is compared (by humans or automatically) with an "ideal" summary. Extrinsic methods are usually domain- or query-dependent, but two popular methods which are relatively generic are *relevance assessment*, where the summarizer acts as the back-end to an information retrieval system, and *reading comprehension*, where the summaries are used as input to a question-answering task. In both cases the idea is to compare performance of the task given the summaries versus the whole documents.

Though it could be argued that reading comprehension is somewhat dependent on coherence, almost all evaluation methods are designed primarily to assess topic coverage and information relevance. This may be because to date, researchers have concentrated on evaluation of highly-compressed summaries, where coherence necessarily takes a back seat to topic coverage. For digesting, even a random selection of sentences is likely to cover all the major topics, so the focus should be on maximizing coherence and minimizing redundant and irrelevant passages.

Another reason why coherence is not measured directly is the dearth of good, automatable evaluation metrics for the trait. One approach commonly used in essay assessment (see overview in Miller [2003]) is to average the semantic similarity (using the cosine coefficient, with or without LSA) of all adjacent sentence pairs. Of course, this technique is not appropriate for our summaries, since by definition of our algorithm, they are guaranteed to have good intersentential cosine scores. This approach has the additional disadvantage of rewarding redundancy.

A more recent approach to automated coherence assessment is to check for the presence or absence of discourse relations [Marcu, 2000]. Since there are no robust programs capable of identifying such relations among arbitrary spans of text, counting unresolved surface-level discourse markers is sometimes employed as a fallback technique (*e.g.,* Nadeau and Tourigny, 2001). The problem with this approach is that the vast majority of discourse relations are not signalled by an obvious discourse marker [Marcu and Echihabi, 2002]. For example, consider the following:

> Cecil likes parrots. Magdeline hates anything with wings.

This sentence pair illustrates a contrast relation, but there is no helpful cue phrase, such as "but" or "however", to indicate this. Only our knowledge of the semantic relations between "likes" and "hates", and "parrots" and "wings", permits us to infer the discourse relation.

Since we also could not come up with a new task-based evaluation which would measure coherence in isolation, we felt we were left with no choice but to use the intrinsic method of quality evaluation. We therefore recruited human judges to provide ratings for our summaries' coherence, and for the sake of convenience and simplicity, we also used them to assess other aspects of summary quality.

3.1.2 Experiment

Source data

We initially considered using the TIPSTER Information-Retrieval Text Research Collections which are used by the annual Document Understanding Conference (DUC), as this might have facilitated comparison and interpretation of our results. However, we found that most of the DUC documents were very short and focussed on single, narrow topics, making them unsuitable for an evaluation of summary coherence. We therefore randomly searched a recently published encyclopedia until we found an article of about 1000 words and another of about 2000 words. We also randomly selected one of the five longest newspaper articles from the DUC 2001 trial data. Our final selections were a 1850-word *Wall Street Journal* article on the 1992 U.S. presidential elections [Murray, 1992], and encyclopedia articles on the English civil war and Kazakhstan [Columbia Encyclopedia, 2001a,b]. The documents, henceforth referred to as `pres92`, `civilwar`, and `kazakhstan`, were stripped of metadata (titles, byline, *etc.*) and entered into a computer as plain text.

Comparison systems

Unfortunately, we were unable to obtain most of the related summarization systems discussed in §1.2. The summarizers we were able to obtain are listed below. Abbreviations referring to these systems (and to the baselines) in our graphs and tables are given in Table 3.1.

Lal and Rüger [2002] have developed an extract-based summarizer built within the GATE[1] framework [Cunningham *et al.*, 2002]. The as-yet nameless system works as a Bayesian classifier over sentences, using features such as sentence and paragraph position,

[1] General Architecture for Text Engineering

word count, and presence of named entities. The system attempts to resolve pronominal anaphora, making its summaries not quite extracts. It also has an optional lexical simplification component, which we disabled for our trial runs. While the authors do not discuss textual coherence in their paper, they do indicate that the purpose of the summarizer is to assist elementary-school students with reading comprehension.

Microsoft Word [Microsoft Corporation, 2002] is a popular commercial word processor which includes an automatic summarization component. We could find no papers describing its inner workings, but it appears to produce extracts rather than abstracts. We feel that it is a valuable basis of comparison because it is so widely used. Other researchers in the field (*e.g.,* Marcu, 1999) have also employed the Word summarizer as a benchmark.

Copernic [Copernic Technologies, 2001, 2002] is a standalone commercial summarization system. While the details provided in the company's white paper are sketchy, we understand that the system employs a Bayesian classifier, a topic segmenter, and numerous language-specific NLP tools. Copernic integrates the National Research Council of Canada's Extractor [Turney, 2000] to identify keyphrases, which are then used to help identify relevant sentences for extraction [Nadeau and Tourigny, 2001].

Sinope [Carp Technologies, 2001] is a commercial version of the Sumatra summarizer (see §1.2) developed at the University of Twente. Sinope was selected for inclusion in this study because, like our summarizer, it attempts to identify latent semantic relations in the text, and uses these relations to determine which text to include in the summary.

Baselines

There are two popular methods for constructing baseline extracts of a given length, both of which are used in our study. The first is to randomly select n sentences from the

Identifier	Summarization system
`copernic`	Copernic
`init`	initial-sentences baseline
`lsa`	our system
`nolsa`	our system minus the LSA component
`plal`	Lal and Rüger, 2002
`random`	random-sentences baseline
`sinope`	Sinope
`word`	Microsoft Word summarizer

Table 3.1: Summarizer name abbreviations

document and present them in their original order of appearance. The second way, based on the observation [Baxendale, 1958] that important sentences are usually located at the beginning of paragraphs, is to select the initial sentence of the first n paragraphs. If the document has fewer than n paragraphs, then the second (and, if necessary, third and subsequent) sentences of the paragraphs are also selected. This approach is commonly referred to as the *lead-* or *initial-sentences baseline*.

In order to measure the contribution of LSA to our system's performance, we also employed a version of our summarizer which does not include the LSA component. Like the base system, it generates a term–sentence co-occurrence matrix and uses it to compute cosine coefficients as a measure of sentence similarity, but the matrix does not undergo singular value decomposition and dimensional reduction. The non-LSA system is in every other respect identical to our base system.

Test procedure

We ran the eight summarizers on the three source documents twice each—once to produce a "short" summary (around 100 words) and once to produce a "long" summary (around 300 words). Most of the summarizers we used do not allow one to specify the maximum summary length in terms of the number of words, but all of them allow specification of the compression ratio as a percentage of the source document length.[2] Accordingly, we used compressions of 15% and 30% for the `kazakhstan` document, and 5% and 15% for the longer `pres92` and `civilwar` documents. Pursuant to our findings in §2.3, our LSA summarizer was set to retain 25% of the singular values for `kazakhstan`, and 20% for the two other documents.

We then recruited human judges by selecting the first 18 volunteers who responded to an advertisement sent to computer science graduate students at the University of Toronto. All the volunteers self-identified as fluent in English, the language of the three source documents. The judges were provided with these documents and the 48 summaries grouped according to source document and summary length. Within each document–summary length group, the summaries were labelled only with a random number and were presented in random order.[3] We asked the judges to read each source document and then assign to each of its summaries an integer score ranging from 1 (very poor) to 5 (very good) on each of three dimensions: comprehensiveness, coherence, and overall quality. The judges were given the compression ratio for each summary and told to take it under consideration when assigning their ratings. To help us interpret the experiment results, the judges were encouraged to explain their ratings and make additional comments in writing.

3.2 Results

3.2.1 Interjudge agreement

To compare interjudge agreement, we computed correlation matrices for each of coherence, comprehensiveness, and overall quality ratings. The results are summarized in the box-and-whisker plots of Figures 3.1 through 3.3, which show the mean interjudge Pear-

[2] Some summarizers interpret the percentage in terms of words, and others in terms of sentences. Given the low variance in sentence length, we do not consider this to be problematic.

[3] The same order was used for each judge.

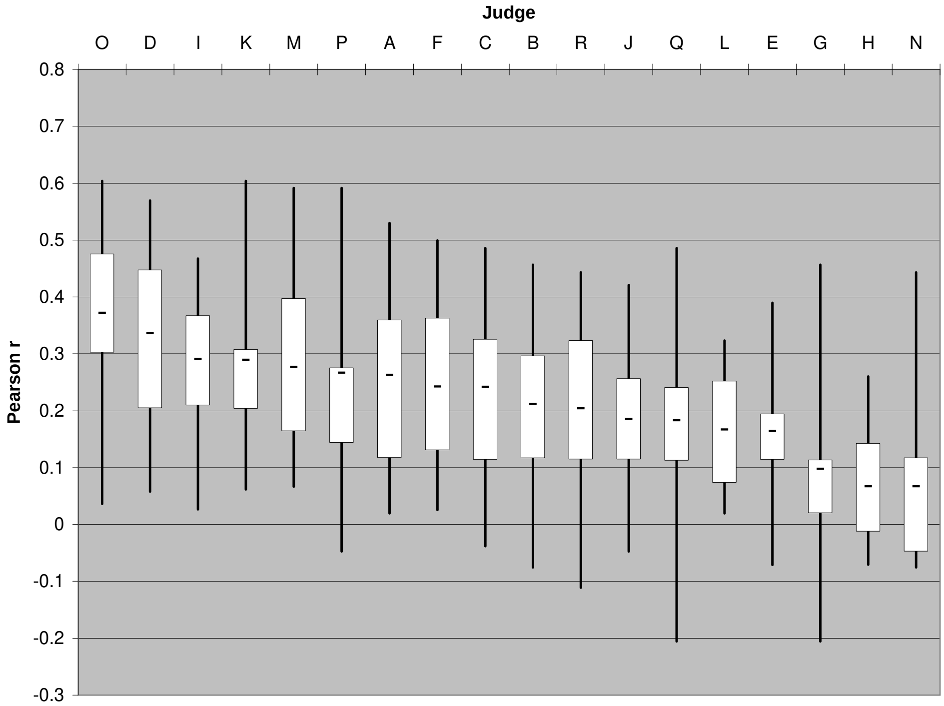

Figure 3.1: Interjudge agreement on coherence

son correlation for each judge. (The convention used in this paper for box-and-whisker plots has the whiskers extending to the minimum and maximum values, the box extending to the first and third quartiles, and the mean value dividing the box.) Interjudge agreement on coherence was generally low, with mean r ranging from 0.0672 to 0.3719. Agreement on comprehensiveness and quality was better, but still only moderate, with r in the ranges $[0.2545, 0.4660]$ and $[0.2250, 0.4726]$, respectively. Why the correlation is only moderate is difficult to explain, though it was not entirely unexpected. Lin and Hovy [2002] report that in the DUC 2001 evaluations, interjudge agreement was around 40% in the single-documents task, and even lower (around 29%) in the multi-documents task.[4]

It is not entirely clear why agreement on coherence, however, should be so low.

[4]The DUC 2001 evaluation method differed from ours in that assessors made pairwise comparisons of system-generated summaries to “ideal” human-generated ones. However, we cite it here for the purpose of demonstrating that instability of manual judgments is not unique to our method.

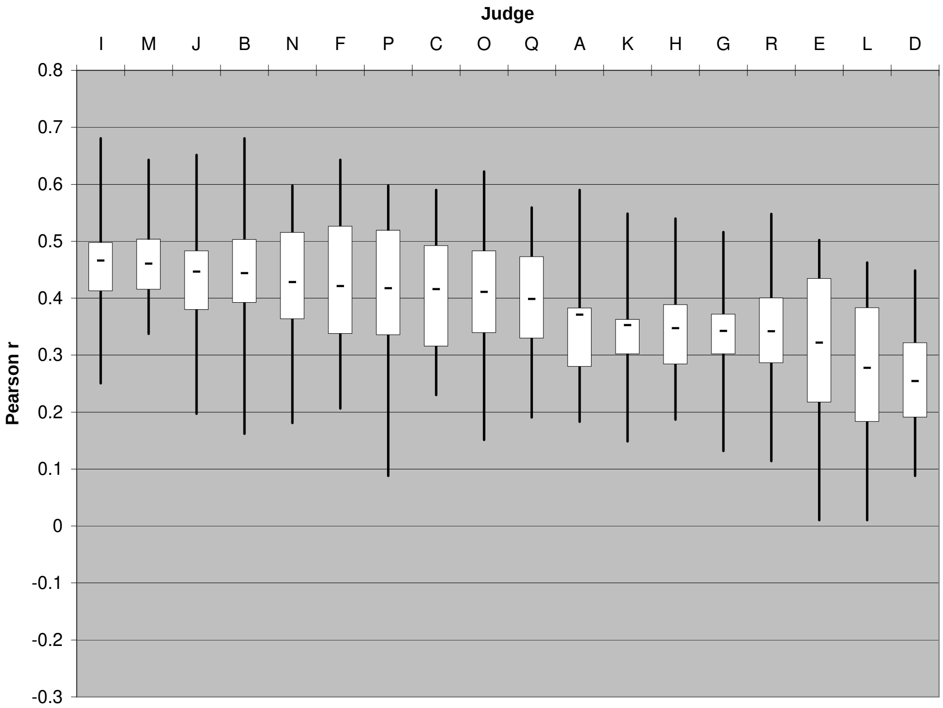

Figure 3.2: Interjudge agreement on comprehensiveness

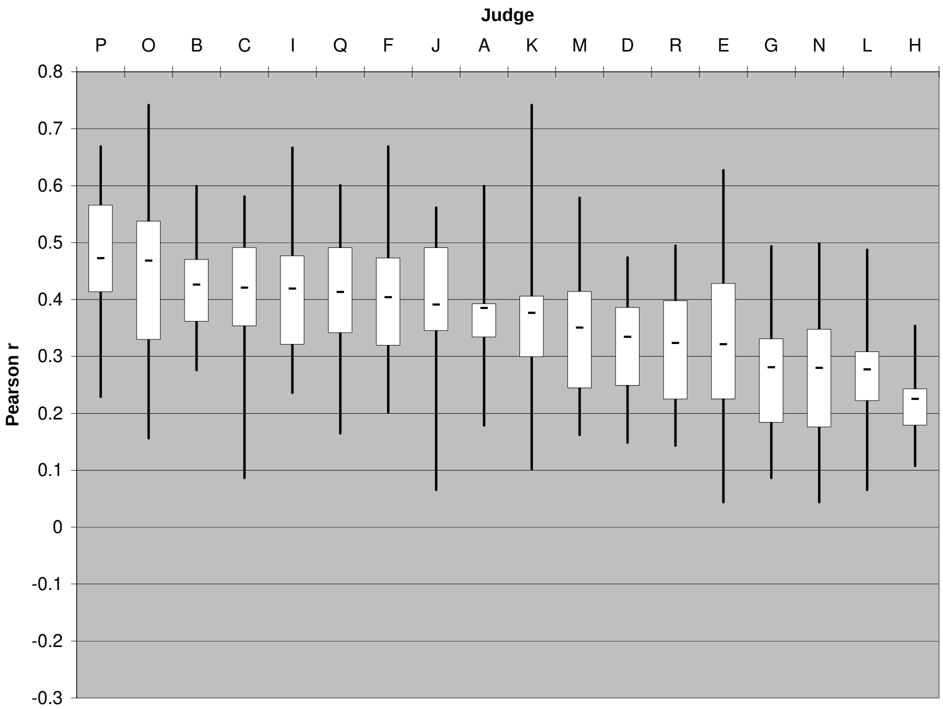

Figure 3.3: Interjudge agreement on overall quality

Though we had made an effort to narrowly define coherence in the written instructions to the judges, it is possible that some of them nevertheless conflated the term with its more conventional meaning of intelligibility, or with cohesion. This last possibility seemed to be reinforced by the judges' written comments, many of which expressed annoyance at dangling discourse markers that actually had little or no bearing on textual coherence as we define it. These cases usually involved uses of "also" or "meanwhile" referring to a previous sentence which the summarizer had not deemed relevant enough to include, and whose exclusion did not appear (at least to us) to constitute a significant breach in topic flow. We suspect that simply removing the marker would have sufficed to allay the judges' complaints.

We therefore considered the possibility that certain groups of judges had interpreted the instructions in different ways. However, attempts to find such groupings were fruitless. Though the most notable outliers—judges G, H, and N—each correlated very poorly with the other judges, they did not correlate well with each other either. Perhaps interjudge agreement is low simply because textual coherence is a very subjective concept.

3.2.2 Comparative performance of summarizers

We used SAS to perform a three-way repeated-measures analysis of variance (ANOVA) for each of the three dimensions: coherence, comprehensiveness, and overall quality. Quite unexpectedly, the (*document*, *summary length*, *summarizer*) three-way interaction effect was significant at the 0.05 confidence level for all three dimensions ($p = 0.0151$, $p < 0.0001$, and $p = 0.0002$, respectively). This means it would have been very difficult, if not impossible, to make any generalizations about the performance of the individual summarizers. On the reasonable[5] assumption that the type of document was irrelevant to summarizer performance, we added the document scores for each (*summarizer*, *summary length*, *rater*) triplet to get new coherence, comprehensiveness, and overall quality mea-

[5]None of the summarizers we tested claimed to be tied to a particular source document type or genre.

surements in the range $[3, 15]$. We then performed two-way repeated-measures ANOVAs for each dimension. The two-way interaction effect was still significant for comprehensiveness ($p = 0.0025$) and overall quality ($p = 0.0347$), but not for coherence ($p = 0.6886$). We now discuss the results for each dimension individually.

Coherence

In our coherence ANOVA, the only significant effect was the summarizer ($p < 0.0001$). That summary length was not found to be significant ($p = 0.0806$) is somewhat surprising, since we expected a strong positive correlation between the coherence score and the compression ratio. Though we did ask our judges to account for the summary length when assigning their scores, we did not think that very short extracts (as opposed to abstracts) could maintain the same level of coherence as their longer counterparts. It may be that summary length's effect on coherence is significant only for summaries with much higher compression ratios than those used in our study.

With respect to the comparative performance of the summaries, only 7 of the 28 pairwise comparisons from our ANOVA were significant at the 0.05 confidence level. The initial-sentences baseline was found to perform significantly better than every other summarizer ($p \leq 0.0008$[6]) except `copernic` and `plal`. The only other significant result we obtained for coherence was that the `sinope` summarizer performed worse than `copernic` ($p = 0.0050$) and `plal` ($p = 0.0005$). Using these pairwise comparisons, we can partition the summarizers into three overlapping ranks as shown in Table 3.2. Further observations on the variance are summarized in the box-and-whisker plot of Figure 3.4.

Comprehensiveness

The mean comprehensiveness score for long summaries was higher than that for short summaries by a statistically significant 1.9792 ($p < 0.0001, \alpha = 0.05$). In fact, in no

[6] All p values in this chapter from here on are Tukey-adjusted.

Rank(s)		Summarizer	Mean rating
A		`init`	11.1111
A	B	`plal`	9.9722
A	B	`copern`	9.6667
C	B	`word`	8.9444
C	B	`lsa`	8.7222
C	B	`nolsa`	8.6667
C	B	`random`	8.4722
C		`sinope`	7.7500

Table 3.2: Summarizer coherence rankings

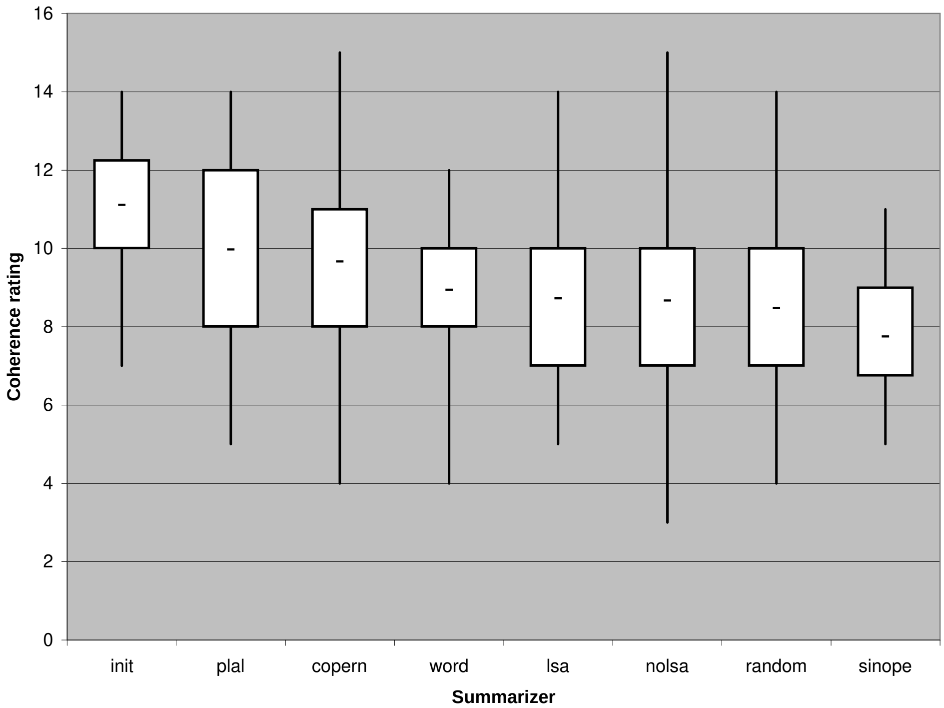

Figure 3.4: Summarizer coherence ratings

case did any summarizer produce a short summary whose mean score exceeded that of the long summary for the same document. This could be because none of the short summaries covered as many topics as our judges thought they could have, or because the judges did not or could not completely account for the compression level.[7] In order to resolve this question, we would probably need to repeat the experiment with abstracts produced by human experts, which presumably have optimal comprehensiveness at any compression ratio.

As with coherence, we can partition the summarizers into overlapping ranks based on their statistically significant comprehensiveness scores. Because the (*summary length*, *summarizer*) interaction was significant, we produce separate rankings for short and long summaries. (See Table 3.3.) Also because of this significance, we expect (and observed) less differentiation among the long summaries, since, as we noted in §3.1.1, simply having more sentences in the extract increases the likelihood of covering more topics.

Short summaries				**Long summaries**			
Rank(s)		Summarizer	Mean rating	Rank(s)		Summarizer	Mean rating
A		`copern`	10.0556	A		`plal`	11.9444
A		`plal`	9.6667	A	B	`copern`	10.5556
A	B	`init`	8.5556	A	B	`init`	10.2222
A	B	`nolsa`	8.1111		B	`sinope`	9.6667
	B	`lsa`	7.5556		B	`word`	9.6111
C	B	`sinope`	7.0000		B	`random`	9.2222
C	B	`word`	6.9444		B	`lsa`	8.9444
C		`random`	5.3889		B	`nolsa`	8.9444

Table 3.3: Summarizer comprehensiveness rankings

The statistics on comprehensiveness ratings are more fully summarized in Figures 3.5 and 3.6.

[7] Only one judge actively demonstrated conscientiousness about the compression ratio, frequently citing in her written comments the "size limitations" of the summary when pointing out areas of redundancy and irrelevancy.

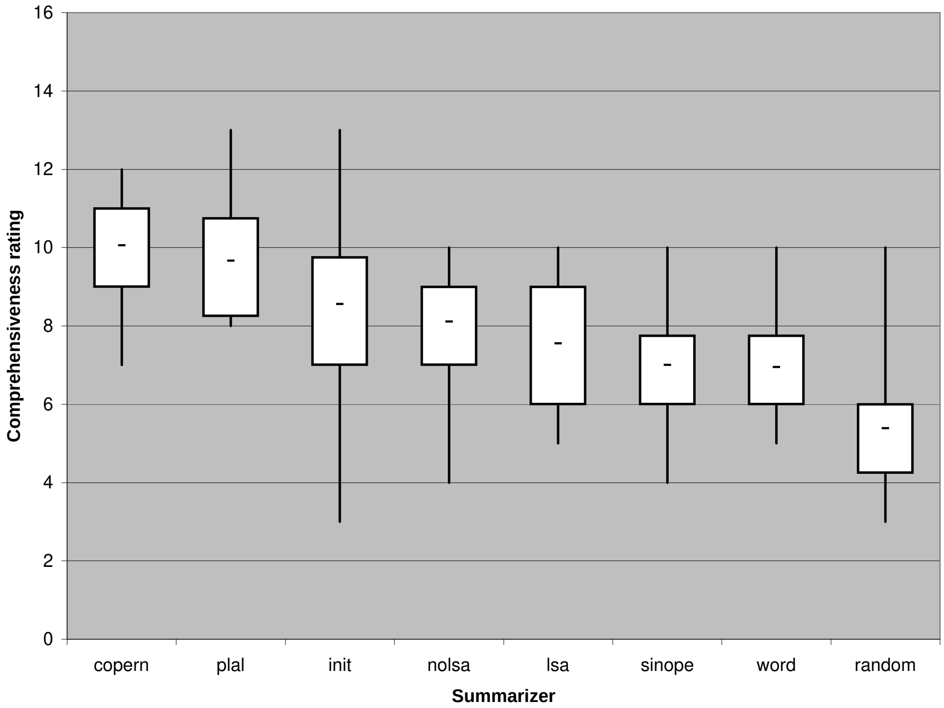

Figure 3.5: Summarizer comprehensiveness ratings (short summaries)

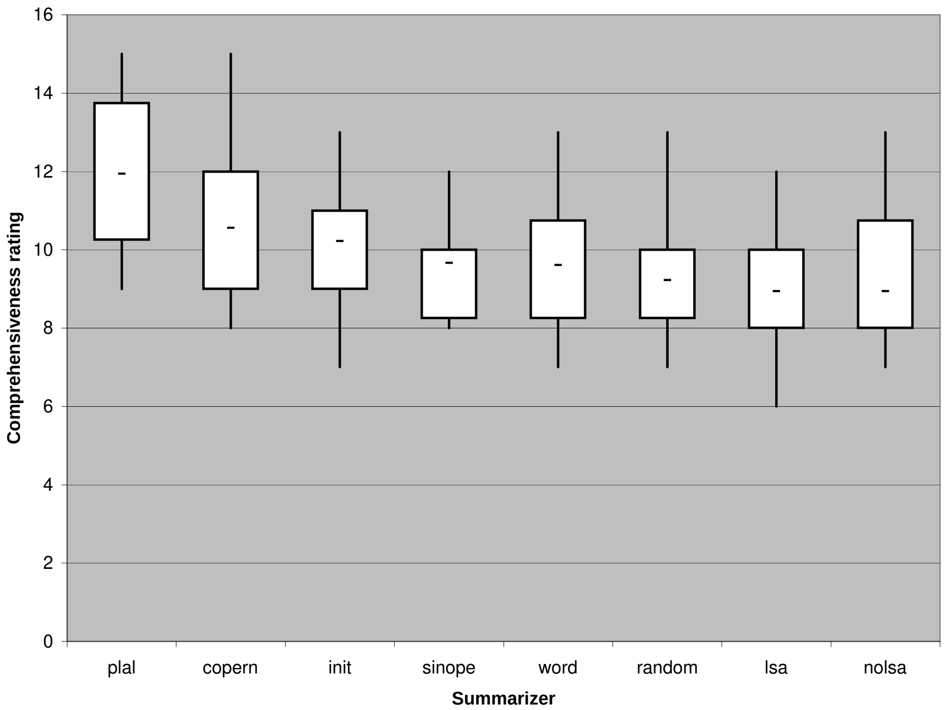

Figure 3.6: Summarizer comprehensiveness ratings (long summaries)

Overall quality

As with comprehensiveness, overall quality scores were dependent not only upon the summarizer but also the summary length. Again, it is not clear whether this is because our judges did not factor in the compression ratio, or because they genuinely believed that the shorter summaries were not as useful as they could have been for their size. Unfortunately, there is nothing in the judges' written comments that sheds light upon what factors they may have considered when assessing overall quality. The rankings and graphs are shown in Table 3.4 and Figures 3.7 and 3.8.

Short summaries				**Long summaries**			
Rank(s)		Summarizer	Mean rating	Rank(s)		Summarizer	Mean rating
A		`copern`	9.7222	A		`plal`	11.1667
A	B	`init`	9.4444	A	B	`init`	10.2778
A	B	`plal`	9.0556	A	B	`copern`	9.9444
A	B	`nolsa`	7.5000	A	B	`word`	9.2222
C	B	`lsa`	7.3333	A	B	`lsa`	9.0556
C		`word`	6.9444		B	`random`	8.5000
C		`sinope`	6.7778		B	`nolsa`	8.3333
C		`random`	5.5556		B	`sinope`	8.1667

Table 3.4: Summarizer overall quality rankings

3.2.3 Relationship among dimensions

Intuition tells us that overall quality of a summary depends in part on both its topic flow and its topic coverage. To see if this assumption is borne out in our data, we calculated the Pearson correlation coefficient for our 864 pairs of coherence–overall quality ratings and comprehensiveness–overall quality ratings. The correlation between coherence and overall quality was strong at $r = 0.6842$, and statistically significant ($t = 27.55$) below the 0.001 confidence level. The comprehensiveness–overall quality correlation was also quite strong ($r = 0.7515, t = 33.44, \alpha < 0.001$).

We expect the relationship between coherence and comprehensiveness to vary with the extract length. For very highly compressed extracts, high coherence can often be

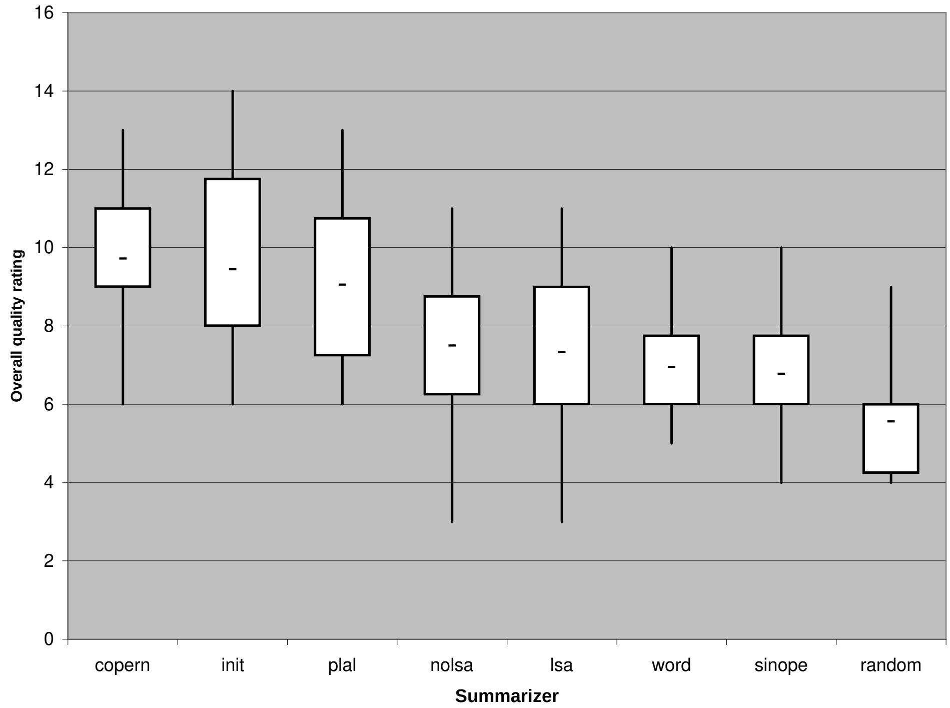

Figure 3.7: Summarizer overall quality ratings (short summaries)

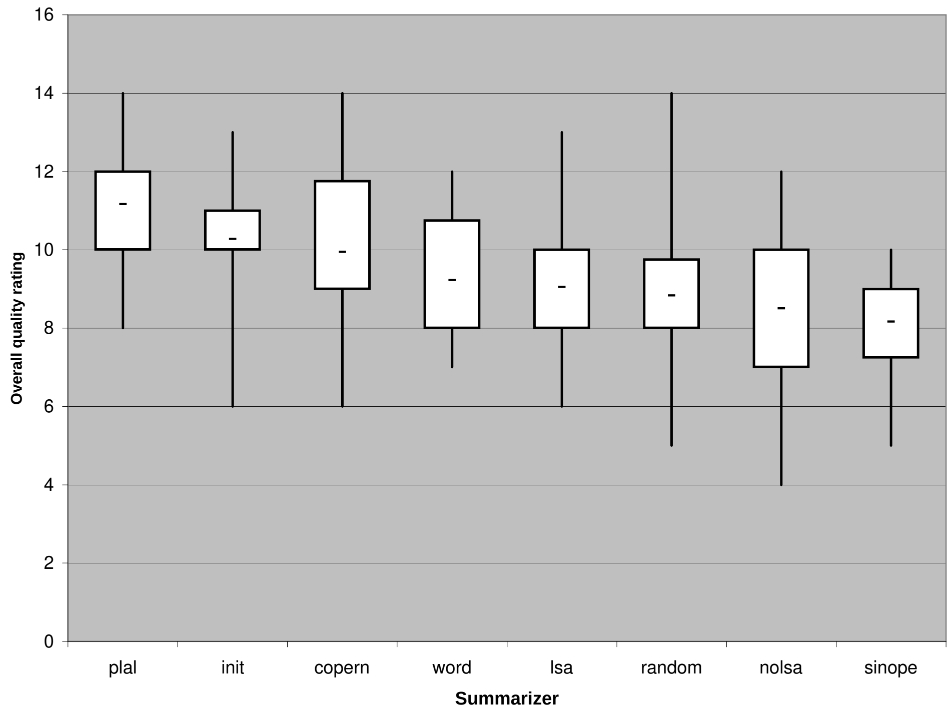

Figure 3.8: Summarizer overall quality ratings (long summaries)

obtained, at the expense of topic coverage, by extracting only sentences which sequentially follow each other in the source document. On the other hand, selecting a minimum amount of material from each topic segment would yield a topically-broad yet very incoherent summary. For summaries with lower compression, however, it should be possible to obtain high topic coverage without compromising textual coherence. Such were the summaries produced in our experiment, which had a coherence–comprehensiveness correlation of 0.4183. We did not observe much of a difference between the short summaries ($r = 0.4200$) and the long summaries ($r = 0.4174$) in this regard.

3.3 Analysis

Unfortunately, moderate to low interjudge agreement for all three dimensions, coupled with an unexpected three-way interaction between the summarizers, the source documents, and the compression ratio, hampered our attempts to make high-level, clear-cut comparisons of summarizer performance. The statistically significant results we did obtain have confirmed what researchers in automatic summarization have known for years: that it is very hard to beat the initial-sentences baseline. This baseline consistently ranked in the top category for every one of the three summary dimensions we studied. While the `copern` and `plal` systems sometimes had higher mean ratings than `init`, the difference was never statistically significant.

Predictably, the random-sentences baseline was characterized by a wide range of scores but very poor performance overall. Like `random`, the `word` and `sinope` systems languished in the bottom rank for every dimension. A glance at the judges' written comments, as well as the summaries themselves, reveals that `sinope` was stymied by an insufficiently discriminating sentence-boundary detection routine. Virtually every occurrence of a period was taken to be a sentence boundary, resulting in strings of incoherent sentence fragments terminated by abbreviations such as "Mr.". This behaviour was almost certainly responsible, at least in part, for the unexpected statistical significance of the (*summarizer*, *document*) interaction effect, since it was only in the `pres92` document that these abbreviations abounded. As for `word`, it had no problems anywhere nearly as glaring, so it is unclear to us why it placed as low as it did.

The performance of our own systems was unremarkable; they consistently placed in the second of the two or three ranks, and only once in the first as well. As with `word`, the judges' notes and summaries do not provide enough information for us to speculate as to why they ranked as they did. It is interesting to note, however, that the judges apparently did not consider segmentation of the summaries into topically oriented paragraphs to be a great benefit. Only our systems and `word` attempted such a segmentation, but none of them appear among the top-ranked for coherence. Only one judge opined that the one-sentence paragraphs emitted by the other summarizers were "hard to read".

Finally, though one of the main foci of our work was to measure the contribution of the LSA metric to our summarizer's performance, we were unable to prove any significant difference between the mean scores for our summarizer and its non-LSA counterpart. The two systems consistently placed in the same rank for every dimension we measured, with mean ratings differing by no more than 6%. Nevertheless, perhaps an informal survey of one `nolsa`–`lsa` summary pair may give us some insight into the differences between these systems.

3.3.1 Case study: `civilwar`

Let us take the example of the long (15%) summaries produced for the `civilwar` document, notable for its length, narrative structure, and range of topics—factors which make summary coherence all the more important. The original document is reproduced in §A.2, with the paragraphs numbered and the topic boundaries (as found by C99) indicated by section markers. Marked-up versions of the summaries are shown here in Figures 3.9 and 3.10, again with topic boundaries marked, and also with the representa-

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans. The 16th cent. had seen a decline in the influence of the nobility and a striking rise in the numbers, wealth, and influence of the gentry and merchants, the beneficiaries of a tremendous expansion of markets and trade in Tudor times. § James had little understanding of the popular unrest and aroused deeper opposition by his continued collection of impositions and benevolences, his dependence on favorites, and his scheme of a Spanish marriage for his son Charles. The Parliament of 1625 granted him the right to collect tonnage and poundage (customs duties) only for a year and not, as was customary, for his entire reign. Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber. § Those imprisoned by the Star Chamber were freed.

Ship money and tonnage and poundage without parliamentary authorization were abolished. § The radicalism of these demands split the parliamentary party and drove many of the moderates to the royalist side. Armed forces (including many peers from the House of Lords and a sizable minority of Commons) gathered about him in the north. § A Scottish army, under Alexander Leslie, 1st earl of Leven, advanced into Yorkshire early in 1644 and gave aid to the parliamentary army in the north. Unable to join Montrose (who was defeated by Leslie in Scotland) and thwarted in his attempts to secure aid from Ireland or the Continent, the king was unable to halt the steady losses of his party and finally was compelled to surrender himself to the Scots, who made him reassuring but vague promises. § Charles I's son Charles II was recognized as king in parts of Ireland and in Scotland but was forced to flee to the Continent after his defeat at Worcester (1651).

Figure 3.9: `nolsa`'s 15% summary of the `civilwar` document

tive sentences from each segment underlined. The first sentence is also underlined, since both algorithms include it by default as an anchor for glue sentences. All sentences not underlined are glue sentences.

Topic sentences

For this document, the C99 topic segmenter partitioned the article into six sections, roughly corresponding to (in order) the background of the struggle, the opposition to James I and Charles I, the initial acts of the Long Parliament, the struggle between Charles I and the Long Parliament, the first civil war, and the second civil war. On the

> English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy. That name, however, overemphasizes the religious element at the expense of the constitutional issues and the underlying social and economic factors. The Parliament that met in 1604 soon clashed with the king on questions of finance and supply. § The Parliament of 1626 went further and impeached the king's favorite, George Villiers, 1st duke of Buckingham. Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber. § Those imprisoned by the Star Chamber were freed. Strafford was impeached, then attainted and executed (1641) for treason; Laud was impeached and imprisoned. § Despite the king's compliance to the will of the opposition thus far, he was not trusted by the parliamentary party. This encouraged Charles to assert himself, and in Jan., 1642, he attempted to arrest in person Pym and four other leaders of the opposition in Commons. § Charles managed to cut off Essex in the southwest but shortly thereafter met parliamentary troops from the north in an indecisive engagement at Newbury. § The legislative remnant known as the Rump Parliament erected a high court of justice, which tried the king for treason and found him guilty. Charles was beheaded on Jan. 30, 1649, and the republic known as the Commonwealth was set up, governed by the Rump Parliament (without the House of Lords) and by an executive council of state.

Figure 3.10: `lsa`'s 15% summary of the `civilwar` document

whole, however, neither `lsa` nor `nolsa` chose particularly intuitive representatives from these topics. The `nolsa` system chose to open with a statement on the rise in power of the mercantile class, which, while relevant, does not in itself explain the fundamental constitutional issue leading to the war. The `lsa` system seems to have better understood the problem by selecting a sentence which makes reference to the war's social and economic roots, but unfortunately it is too vague to be useful. Perhaps the last sentence of ¶2 in the original document, which clearly sets out the quarrel between Parliament and the king, would have served better. The representatives chosen from the second and third topic segments are also unhelpful, relating specific acts Parliament made against Charles I rather than raising the general issue of the kings' despotism. Here the second sentences of ¶7 and ¶11 better embody the hearts of their respective topic segments; it is not clear why sentences such as these were not chosen instead.

For the next topic segment, `lsa` makes a good selection with a sentence establishing Parliament's mistrust of Charles I. The `nolsa` system also fares well here in that it explains how the bipartisanship within the ruling class developed, though some anaphora resolution (such as replacing "these demands" with "Parliament's demands for reform") would have made things clearer. Neither system chooses a particularly salient sentence from the first civil war topic, though a glance at this section in the original document reveals that there are really no sentences which could be considered truly representative. For the last topic, `lsa` again does well in relating the execution of the king and the establishment of the Commonwealth, which were the major outcomes of the war. By contrast, a relatively unimportant sentence is selected by `nolsa`.

Glue sentences

Turning now to the glue sentences, it seems to us that `nolsa`'s choices are much more shallow and transparent than its counterpart's, relying only on keywords common to either or both anchor sentences. The first glue sentence seems to have been chosen on the basis of the occurrence of "gentry" and "merchants" in the latter anchor. Nevertheless, it does make for a good segue between the first and third sentences of the summary. The next glue sentence is unusual in that it bears little resemblance to either of its anchors, but coincidentally it is one of the sentences we recommended as a good topic sentence for this segment. The third glue sentence seemingly hinges upon the presence of "Parliament" in the first anchor and "Star Chamber" in the second, but fits better with the second anchor than with the first. Like the second glue sentence, the fourth glue sentence is lexically dissimilar to either of its anchors, but happens to follow logically from the preceding two sentences. The remaining two glue sentences again hinge on the presence of a few points of commonality with the latter anchor (*e.g.,* "north", "Scotland", "Ireland"), and again these sentences seem to function more as a prelude to the latter topic sentence than as a true bridge between both.

With respect to `lsa`'s performance, its first glue sentence is a perfect match, but only because it is the only sentence to be found between the two anchors in the original document. The second glue sentence shares few words with either of its anchors, but was likely selected on the basis of the high LSA cosine scores between "Parliament" and "issues" ($r = 0.581$), "king" and "social" ($r = 0.592$), and "questions" and "constitutional" ($r = 0.781$) with the first anchor, and "king" and "Parliament" ($r = 0.658$), "king" and "impeached" ($r = 0.637$), and "king" and "king's" ($r = 0.562$) with the second anchor. It serves as a good bridge between the topic sentences because it introduces the conflict between the king and the Parliament, making reference to the economic factors of the first topic sentence and paving the way for the second topic sentence to address "further" acts of Parliament. That the glue sentence and the second topic sentence refer to different kings is forgivable in this case, because it was both James I and Charles I who incited the war, and because neither sentence mentions its respective king by name.

The glue between the next two topic sentences is probably one of the best examples of LSA at work. It seems to have been chosen not just from the co-occurrence of "Parliament" and "Star Chamber", but also on the basis of the high cosine score between "king's" and "opponents" ($r = 0.972$), and between "imprisoned" and both "Star" and "Chamber" ($r = 0.744$ for each). In fact, LSA seems to have learned the semantic association between many words in the article relating to the legal system: "Star", "Chamber", "courts", "imprisoned", "impeached", and "treason" all have high cosine scores with each other. The net effect in this case is that the `lsa` system has selected a sentence which smoothly continues the list of Parliament's acts against the king, while at the same time provides a crucial piece of information necessary to understand the release of the Star Chamber convicts mentioned in the next sentence. This is all the more impressive considering that in the original document, the latter topic sentence occurs several paragraphs after the glue sentence.

The fourth glue sentence does not seem particularly appropriate, but this can be explained by the proximity of the anchor sentences in the original document. Among five potential glue sentences, none of them contain especially important linking material. The search for the fifth glue sentence was undoubtedly hampered by the wholly inappropriate representative sentence chosen for the fifth topic. Perhaps only by accident, the system chose as glue the sentence which best encapsulates the immediate cause of the civil war, even if it does not explicitly label it as such. The final glue sentence, suffering from the same irrelevant anchor, nonetheless manages to unite its two topic sentences using the relationship between "parliamentary" and "Parliament" ($r = 0.457$), "Charles" and "Parliament" ($r = 0.670$), "Charles" and "king" ($r = 0.490$), and the co-occurence of "Rump Parliament". The glue fails to explain how the war was won, but is still very useful in that it gives the reason for the king's later-mentioned execution.

Conclusion

Though we have shown how the use of LSA can sometimes be of benefit in selecting linking material, our human judges did not agree that the LSA-based system we used produced more coherent summaries for this document. The mean coherence scores for these summaries were 3.2778 for `nolsa`, but only 3.0556 for `lsa`. Likewise, `lsa` failed to excel in comprehensiveness (3.2778 *vs.* 3.4444), and was only marginally better when considering overall quality (3.1667 *vs.* 3.1111). Perhaps these results can be explained, at least in part, by the summarizers' poor choice of topic sentences. Without important, informative sentences around which to structure the summary, the question of what constitutes appropriate linking material between them becomes moot. Further study is necessary to determine whether the poor choice of topic sentences was an unfortunate coincidence, or whether we should investigate an entirely new approach to topic sentence extraction. It may be that lexical semantic relationships alone are not sufficient for assessing salience, and that it is also necessary to consider features relating to syntax, pragmatics, and discourse structure.

Chapter 4

Conclusion

4.1 Summary

Our goal in this work has been to investigate how we can improve the coherence of automatically-produced extracts. We developed and implemented an algorithm which builds an initial extract composed solely of topic sentences, and then fills in the lacunae by providing linking material between semantically dissimilar sentences. In contrast with much of the previous work we reviewed, our system was designed to minimize reliance on language-specific features.

Our summarizer differs in architecture from most others in that it measures semantic similarity with latent semantic analysis, a factor analysis technique which builds upon the vector-space model typically used in IR. We believed that the deep semantic relations discovered by LSA would assist in the identification and correction of abrupt topic shifts in the summaries. In order to determine whether LSA had any advantages over the plain cosine similarity metric, we tested our system both with and without the LSA component activated.

An experiment was conducted wherein human judges reviewed summaries produced by our system, its non-LSA counterpart, a summarizer representing the state of the art in research, three summarizers representing the state of the art in commercial systems, and two simplistic baseline systems. For each summary, the judges were asked to provide numerical ratings for coherence (topic flow and organization) and comprehensiveness (topic coverage), as well as a score representing their opinion on the overall quality and usefulness of the summary.

The study provided few clearly-defined distinctions among the summarization systems. Though our evaluation method for coherence was intended to circumvent the limitations of automated approaches, the use of human judges introduced its own set of problems, foremost of which was the low interjudge agreement on what constitutes a fluent summary. Despite this lack of consensus, we found a strong positive correlation between the judges' scores for coherence and overall summary quality. We would like to take this as good evidence that the production of coherent summaries is an important research area within automatic summarization. However, it may be that humans simply find it too difficult to evaluate coherence in isolation, and end up using other aspects of summary quality as a proxy measure.

4.2 Future work

If there is one benefit to the nebulous results we obtained, it is that we now have the opportunity of determining how we could revise our experimental and evaluation methodologies to avoid this situation in future research. In this section, we discuss some of these revisions, as well as some test parameters we may want to vary in future experiments.

4.2.1 Evaluation methodology

As we noted in §3.2.1, low interjudge agreement on coherence may have arisen from confusion of the term "coherence" with the concepts of cohesion and overall intelligibility. Though the judges were provided with written instructions explaining the concepts, it may be better in the future to provide detailed examples of coherent and incoherent sentences. In particular, we may want to include an example such as the following (from Morris and Hirst [1991]), which illustrates how it is possible to have cohesive ties in the text without having a coherent document.

> Wash and core six apples. Use them to cut out the material for your new suit. They tend to add a lot to the color and texture of clothing. Actually, maybe you should use five of them instead of six, since they are quite large.

We might also provide some examples of coherent texts which use nonsense words (*e.g.,* Lewis Carroll's "Jabberwocky"); this might help dispell the notion that coherence is the same thing as intelligibility.

External factors may also be contaminating the coherence scores. One thing we might do to minimize this risk is to present the summaries to the judges *before* they read the source documents. Since knowledge of the original document is not necessary to gauge the topic flow and organization of the summary, the judges' opinions would not be biased by cases where some crucial piece of information from the source document is omitted or misrepresented.

Even if the issue with interjudge agreement were resolved completely, however, the use of human judges is still problematic. Even when monetary compensation is offered, it is difficult to find volunteers willing to spend the many hours necessary to read through large piles of documents and summaries. It has been suggested to us that this problem could be partly alleviated by using a fractional factorial experiment [Montgomery, 2000, p. 303]; this would allow us to compare a greater number of document sets while at the same time reducing the time commitment of the individual judges.

A better solution might involve abandoning human judges altogether in favour of a fully automated technique, but as we discussed in §3.1.1, all such existing techniques are unacceptable for various reasons. Recent work by Marcu and Echihabi [2002] may be changing this, however. They have developed an unsupervised machine learning system for recognizing discourse relations that hold between arbitrary spans of text. The system is successful (up to 93% accuracy) with some types of discourse relations even when there is no cue phrase explicitly marking the relation. The authors feel that their technique could be adapted to identify a wider range of discourse relations, and even to develop a new, empirically justified classification scheme for discourse relations.

4.2.2 Experimental parameters

In the experiment presented in this paper, we varied only one of the many parameters for our summarizer—namely, the use of LSA on the term–sentence co-occurrence matrix. A fully automated evaluation methodology would make it easier to measure the effects of the other parameters for our algorithm, and to arrive at an optimal configuration. For instance, our informal pre-experiment trials of the system led us to believe that a dimensional reduction of 70–80% produced good summaries. In light of the judges' assessments, however, it is possible we were mistaken in this regard. Changing the dimensional reduction may also necessitate alteration of the α, β, and γ cutoffs; a learning algorithm might help determine the best combination of values.

In future experiments we may also wish to investigate summarizer performance on other document types. In this thesis, we used only short expository texts. Coherence plays a quite different, arguably more important, role in narrative texts and in spoken dialogue. These types of documents place a greater emphasis on temporal relations and question–answer pairs, many of which have sizeable intervening gaps. Are the latent semantic relations embodied in these structures strong enough to help our algorithm pair causes with their effects, and questions with their answers?

Future experiments may also investigate digests of much longer expository texts. In digesting, the process of summarization can be seen more as deciding what irrelevant information to throw away rather than what relevant information to keep. When a large block of text is omitted from a source document, the resulting lacuna may leave a jarring gap in the rhetorical structure of the text. How effective is our algorithm at bridging these sorts of gaps? Maintaining textual coherence is important in these situations, since unlike with today's highly compressed computer-generated summaries, the reader of a digest expects the text to flow nearly as smoothly as in the original.

Appendix A

Selected source documents

A.1 kazakhstan

1 Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. It borders on Siberian Russia in the north, China in the east, Kyrgyzstan, Uzbekistan, Turkmenistan, and the Aral Sea in the south, and the Caspian Sea and European Russia in the west. Astana is the capital and Almaty (Alma-Ata) is the largest city. Other major cities include Shymkent, Semey, Aqtobe, and Oskemen.

2 Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast. It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts. in the east. It is largely lowland in the north and west (W Siberian, Caspian, and Turan lowlands), hilly in the center (Kazakh Hills), and mountainous in the south and east (Tian Shan and Altai ranges). Kazakhstan is a region of inland drainage; the Syr Darya, the Ili, the Chu, and other rivers drain into the Aral Sea and Lake Balkash. Most of the region is desert or has limited and irregular rainfall.

3 The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars. Kazakh, a Turkic language, is the official tongue, but Russian is still widely used. There is considerable friction between the now dominant Kazakhs and the formerly favored ethnic Russians, who continue to emigrate in large numbers. Almaty is the site of Kazakhstan Univ. (founded 1934) and the Kazakh Academy of Sciences (founded 1946).

4 Despite Kazakhstan's largely arid conditions, its vast steppes accommodate both livestock and grain production. In the 1950s, the Virgin Lands Program under Khrushchev brought hundreds of thousands of Russian, Ukrainian, and German settlers to the area. Wheat, cotton, sugar beets, and tobacco are the main crops. The raising of cattle and sheep is also important, and Kazakhstan produces much wool and meat. In addition, there are rich fishing grounds, famous for their caviar-producing sturgeon, in the N Caspian.

5 The Kazakh Hills in the core of the region have important mineral resources. Coal is mined at Qaraghandy and Ekibastuz, and there are major oil fields in the Emba basin (which includes the important Tengiz fields), at the northeast tip of the Caspian Sea, and in the Mangyshlak Peninsula. Kashagan, a Caspian field that was being explored in the late 1990s, appears to have great potential. A pipeline was built in the 1990s to connect the nation's oil fields to the Black Sea. Kazakhstan also has large deposits of natural gas, iron ore, manganese, chrome, lead, zinc, silver, copper, nickel, titanium, bauxite, and gold. The Irtysh River hydroelectric stations are a major source of power.

6 The country's industries are located along the margins of the country. Steel, agricultural and mining machinery, superphosphate fertilizers, phosphorus acids, artificial fibers, synthetic rubber, textiles, and medicines are among the manufactured goods. Temirtau is the iron and steel center. Semey was the Soviet center of space-related industries, and the surrounding region was the site of Soviet nuclear testing; radiation pollution is widespread in the area, which experienced a severe economic downturn following the end of nuclear testing in 1991. The Baikonur (Bayqongyr) Cosmodrome in central Kazakhstan was the Soviet space-operations center and continues to serve Russian space exploration through an agreement between the two nations. The main trading partners are Russia, Ukraine, and Uzbekistan.

7 Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote. There is a bicameral parliament, most of whose members are elected, but its powers are limited. The country is divided into 14 administrative units, or oblasts.

8 The original nomadic Turkic tribes inhabiting the region had a culture that featured the Central Asian epics, ritual songs, and legends. These Kazakh groups were conquered by the Mongols in the 13th cent. and ruled by various khanates until the Russian conquest (1730–1840). The 19th cent. saw the growth of the Kazakh intelligentsia. A written literature strongly influenced by Russian culture was then developed.

9 In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army. Organized as the Kirghiz Autonomous SSR in 1920, it was renamed the Kazakh Autonomous SSR in 1925 and became a constituent republic in 1936. During the Stalin era, collectivization was instituted and millions of Kazakhs were forced to resettle in the region's south in order to strengthen Russian rule. In the early 1960s parts of republic saw extensive agricultural development as the Virgin Lands Territory.

10 Kazakhstan declared its independence from the Soviet Union on Dec. 16, 1991. Nursultan Nazarbayev became the country's first president and soon began a gradual movement toward privatization of the economy. In 1994, Kazakhstan signed a series of security agreements with the United States, in which the latter would take control of enriched uranium usable for nuclear weapons and aid Kazakhstan in removing extant nuclear weapons, closing missile silos, converting biological-weapons-production centers, and destroying its nuclear test ranges. These projects were financed by the United States, and many had been completed by late 1999.

11 Elections in 1994 gave a parliamentary majority to allies of Nazarbayev, but they resisted his reform plans. In Apr., 1995, after the 1994 election results were dismissed as

invalid by the constitutional court, he suspended parliament and ruled by decree. New elections in Dec., 1995, gave his allies a majority in parliament but were criticized by the opposition and others as flawed. On the basis of referendums held in 1995 and 1996 that were denounced by the opposition, Nazarbayev's term in office was extended to the year 2000 and his powers were increased. In an election rescheduled to Jan., 1999, Nazarbayev was reelected after disqualifying the major opposition candidate.

12 In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia. In 1997 the capital was moved from Almaty to the more centrally located Astana (formerly Aqmola). In 1999, as Kazakhstan's economy worsened, the government agreed to sell some of its stake in the vast Tengiz oil field. Kazakhstan is a member of the Commonwealth of Independent States.

A.2 civilwar

1 English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth.

2 The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy. That name, however, overemphasizes the religious element at the expense of the constitutional issues and the underlying social and economic factors. Most simply stated, the constitutional issue was one between a king who claimed to rule by divine right and a Parliament that professed itself to have rights and privileges independent of the crown and that ultimately, by its actions, claimed real sovereignty.

3 Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans. The 16th cent. had seen a decline in the influence of the nobility and a striking rise in the numbers, wealth, and influence of the gentry and merchants, the beneficiaries of a tremendous expansion of markets and trade in Tudor times. It was from this middle class of gentry and merchants that the opposition to the crown drew most of its members. Their ambition to do away with financial and commercial restrictions and their desire to have a say in such matters as religious and foreign policies had been severely restrained by the Tudors, but on the accession (1603) of a Scottish king to the English throne the popular party began to organize its strength.

4 James I was not long in gaining a personal unpopularity that helped to strengthen Parliament's hand. At the Hampton Court Conference (1604) he resolutely refused to compromise with Puritans on religious questions. The Parliament that met in 1604 soon clashed with the king on questions of finance and supply. § James was forced to temporize because of his urgent need of money, but the dissolution of the Parliament in 1610 left feelings of bitterness on both sides.

5 A new Parliament met in 1614, and the Commons engaged in quarrels not only with the king but also with the House of Lords. Because it passed not a single statute, this was called the Addled Parliament. James had little understanding of the popular unrest and aroused deeper opposition by his continued collection of impositions and benevolences, his dependence on favorites, and his scheme of a Spanish marriage for his son Charles.

6 Meanwhile a legal battle was being waged in the courts, with Sir Francis Bacon zealously upholding the royal prerogative and Sir Edward Coke defending the supremacy of common law. The king dismissed Coke from the bench in 1616, but the Parliament of 1621 impeached Bacon. The last Parliament (1624) of the reign accompanied its grant of money with specific directions for its use. James's reign had raised certain fundamental questions concerning the privileges of Parliament, claimed by that body as their legal right and regarded by James as a special grant from the crown.

7 Charles I, married to a French Roman Catholic princess, Henrietta Maria, proved more intractable and even less acceptable to the Puritan taste than his father, and Parliament became even more uncompromising in the new reign. The leaders of the parliamentary party—Coke, John Pym, Sir John Eliot, and John Selden—sought ways to limit the powers of the king. The Parliament of 1625 granted him the right to collect tonnage and poundage (customs duties) only for a year and not, as was customary, for his entire reign. The Parliament of 1626 went further and impeached the king's favorite, George Villiers, 1st duke of Buckingham. Charles dissolved it in anger.

8 Failing to raise money without Parliament, he was forced to call a new one in 1628. The new Parliament drew up the Petition of Right, and Charles accepted it in order to get his subsidy. He continued to levy customs duties, an act that the parliamentarians declared illegal under the Petition of Right. Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber. The religious issue also came up, and Commons resisted the king's order to adjourn by forcing the speaker to remain in his chair while Eliot presented resolutions against "popery" and unauthorized taxation.

9 In the succeeding 11 years Charles attempted to rule without a Parliament, resorting to such expedients as ship money (a tax levied originally on seaports but extended by Charles to the entire country) to raise revenue. The reprisals against Eliot and the prosecution of William Prynne and John Hampden aroused widespread indignation. Charles's chief advisers, Archbishop William Laud and Thomas Wentworth, later 1st earl of Strafford, were cordially detested.

10 The ominous peace was broken by troubles in Scotland, where efforts to enforce Anglican episcopal policy led to the violent opposition of the Covenanters and to war in 1639 (see Bishops' Wars) and compelled Charles to seek the financial aid of Parliament. The resulting Short Parliament (1640) once more met the king's request for supply by a demand for redress of grievance. Charles offered to abandon ship money exactions, but the opposition wished to discuss more fundamental issues, and the king dissolved the Parliament in just three weeks.

11 The disasters of the second Scottish war compelled a virtual surrender by the king to the opposition, and the Long Parliament was summoned (Nov., 1640). § The parliamentarians quickly enacted a series of measures designed to sweep away what they regarded as the encroachments of despotic monarchy. Those imprisoned by the Star Chamber were

freed. A Triennial Act provided that no more than three years should elapse between sessions of Parliament, while another act prohibited the dissolution of Parliament without its own consent. Ship money and tonnage and poundage without parliamentary authorization were abolished. Strafford was impeached, then attainted and executed (1641) for treason; Laud was impeached and imprisoned. Star Chamber and other prerogative and episcopal courts were swept away. However, discussions on church reform along Puritan lines produced considerable disagreement, especially between the Commons and Lords.

12 § Despite the king's compliance to the will of the opposition thus far, he was not trusted by the parliamentary party. This distrust was given sharp focus by the outbreak (Oct., 1641) of a rebellion against English rule in Ireland; an army was needed to suppress the rebellion, but the parliamentarians feared that the king might use it against them. Led by John Pym, Parliament adopted the Grand Remonstrance, reciting the evils of Charles's reign and demanding church reform and parliamentary control over the army and over the appointment of royal ministers. The radicalism of these demands split the parliamentary party and drove many of the moderates to the royalist side. This encouraged Charles to assert himself, and in Jan., 1642, he attempted to arrest in person Pym and four other leaders of the opposition in Commons. His action made civil war inevitable.

13 In the lull that followed, both Parliament and the king sought to secure fortresses, arsenals, and popular support. In June, 1642, Parliament sent to the king a statement reiterating the demands of the Grand Remonstrance, but since the proposals amounted to a complete surrender of sovereignty by the crown to Parliament, the king did not even consider them as a basis for discussion. Armed forces (including many peers from the House of Lords and a sizable minority of Commons) gathered about him in the north. Parliament organized its own army and appointed Robert Devereux, 3d earl of Essex, to head it. On Aug. 22, 1642, Charles raised his standard at Nottingham.

14 The followers of king and Parliament did not represent two absolutely distinct social groups, as the popular conception of the royalist Cavaliers and the parliamentary Roundheads would indicate. However, it is true that the parliamentary, or Puritan, group drew much of its strength from the gentry and from the merchant classes and artisans of London, Norwich, Hull, Plymouth, and Gloucester; it centered in the southeastern counties and had control of the fleet. The majority of the great nobles followed the king, who had the support of most Anglicans and Roman Catholics; geographically the royalist strength centered in the north and west.

15 The first major engagement of the armies at Edgehill (Oct. 23, 1642) was a drawn battle. Charles then established himself at Oxford. § The royalist forces gained ground in the north and west, although repeated attempts by the king to advance on London proved abortive. The indecisive engagements of 1643 were remarkable mainly for the emergence of Oliver Cromwell, an inconspicuous member of the Long Parliament, to military prominence with his own regiment of "godly" men, soon to become famous as the Ironsides.

16 Futile negotiations for peace had been conducted at Oxford early in 1643, and in Sept., 1643, Parliament took a decisive step by securing the alliance of the Presbyterian Scots in accepting the Solemn League and Covenant. Scottish aid was obtained only by a promise to submit England to Presbyterianism, which was soon to produce a reaction

from the Independents and other sectarians (particularly in the army) who opposed the idea of any centralized national church.

17 The war now entered a new phase. A Scottish army, under Alexander Leslie, 1st earl of Leven, advanced into Yorkshire early in 1644 and gave aid to the parliamentary army in the north. Charles's nephew, the brilliant and dashing Prince Rupert, did something to stem royalist losses by retaking Newark, but his gains were temporary. His campaign to relieve the besieged York led to the battle of Marston Moor (July 2, 1644), in which Cromwell and Leslie inflicted a crushing defeat on the royalists. Charles managed to cut off Essex in the southwest but shortly thereafter met parliamentary troops from the north in an indecisive engagement at Newbury.

18 To stem the rising dissension among parliamentary leaders, Cromwell sponsored in Parliament the Self-Denying Ordinance, by which all members of Parliament were compelled to resign their commands, and the parliamentary army was reorganized (1644–45) into the New Model Army. Thomas Fairfax (later 3d Baron Fairfax of Cameron) became the commander in chief.

19 After further futile peace negotiations at Uxbridge, Charles, hoping to join the forces under James Graham, marquess of Montrose, moved north and stormed Leicester. He met Cromwell in a sharp battle at Naseby (June 14, 1645). This battle cost the king a large part of his army and rendered the royalist cause hopeless. Unable to join Montrose (who was defeated by Leslie in Scotland) and thwarted in his attempts to secure aid from Ireland or the Continent, the king was unable to halt the steady losses of his party and finally was compelled to surrender himself to the Scots, who made him reassuring but vague promises. The first civil war came to an end when Oxford surrendered in June, 1646.

20 The king was delivered (1647) by the Scots into the hands of Parliament, but the Presbyterian rule in that body had thoroughly alienated the army. The army resisted Parliament's proposal to disband it by capturing the king from the parliamentary party and marching on London. Army discontent gradually became more radical (see Levelers), and the desire grew to dispose of the king altogether.

21 Refusing to accept the army council's proposals for peace (the Heads of the Proposals), Charles escaped in Nov., 1647, and took refuge on the Isle of Wight, where he negotiated simultaneously with Parliament and the Scots. § In Dec., 1647, he concluded an agreement with the Scots known as the Engagement, by which he agreed to accept Presbyterianism in return for military support. In the spring of 1648, the second civil war began. Uprisings in Wales, Kent, and Essex were all suppressed by the parliamentary forces, and Cromwell defeated the Scots at Preston (Aug. 17, 1648). Charles's hopes of aid from France or Ireland proved vain, and the war was quickly over.

22 Parliament again tried to reach some agreement with the king, but the army, now completely under Cromwell's domination, disposed of its enemies in Parliament by Pride's Purge (Dec., 1648; see under Pride, Thomas). The legislative remnant known as the Rump Parliament erected a high court of justice, which tried the king for treason and found him guilty. Charles was beheaded on Jan. 30, 1649, and the republic known as the Commonwealth was set up, governed by the Rump Parliament (without the House of Lords) and by an executive council of state.

23 Charles I's son Charles II was recognized as king in parts of Ireland and in Scotland

but was forced to flee to the Continent after his defeat at Worcester (1651). The years of the interregnum, under the Commonwealth to 1653 and the Protectorate after that, are largely the story of Oliver Cromwell's personal rule, which was marked by strict military administration and enforcement of the Puritan moral code. After his death and the short-lived rule of his son, Richard Cromwell, the Commonwealth was revived for a brief and chaotic period. It ended in 1660 with the Restoration of Charles II. Although some of the changes brought about by the war were swept away (*e.g.,* in the restoration of Anglicanism as the state church), the settlement of the contest between the king and Parliament was permanently assured in the Glorious Revolution of 1688.

Appendix B

Selected summaries

B.1 kazakhstan

B.1.1 Short summaries (15%)

copernic

It borders on Siberian Russia in the north, China in the east, Kyrgyzstan, Uzbekistan, Turkmenistan, and the Aral Sea in the south, and the Caspian Sea and European Russia in the west.

It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts.

The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars.

Coal is mined at Qaraghandy and Ekibastuz, and there are major oil fields in the Emba basin (which includes the important Tengiz fields), at the northeast tip of the Caspian Sea, and in the Mangyshlak Peninsula.

Semey was the Soviet center of space-related industries, and the surrounding region was the site of Soviet nuclear testing; radiation pollution is widespread in the area, which experienced a severe economic downturn following the end of nuclear testing in 1991.

Nursultan Nazarbayev became the country's first president and soon began a gradual movement toward privatization of the economy.

In Apr., 1995, after the 1994 election results were dismissed as invalid by the constitutional court, he suspended parliament and ruled by decree.

New elections in Dec., 1995, gave his allies a majority in parliament but were criticized by the opposition and others as flawed.

lsa

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. Almaty is the site of Kazakhstan Univ. (founded 1934) and the Kazakh Academy of Sciences (founded 1946). The raising of cattle and sheep is also important, and Kazakhstan produces much wool and meat. Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote. In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army. During the Stalin era, collectivization was instituted and millions of Kazakhs were forced to resettle in the region's south in order to strengthen Russian rule. In the early 1960s parts of republic saw extensive agricultural development as the Virgin Lands Territory. Nursultan Nazarbayev became the country's first president and soon began a gradual movement toward privatization of the economy.

initial

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast. The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars. Despite Kazakhstan's largely arid conditions, its vast steppes accommodate both livestock and grain production. The Kazakh Hills in the core of the region have important mineral resources. The country's industries are located along the margins of the country. Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote. The original nomadic Turkic tribes inhabiting the region had a culture that featured the Central Asian epics, ritual songs, and legends.

nolsa

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts. in the east. Coal is mined at Qaraghandy and Ekibastuz, and there are major oil fields in the Emba basin (which includes the important Tengiz fields), at the northeast tip of the Caspian Sea, and in the Mangyshlak Peninsula. A pipeline was built in the 1990s to connect the nation's oil fields to the Black Sea.

A written literature strongly influenced by Russian culture was then developed. In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army. In Apr., 1995, after the 1994 election results were dismissed as invalid by the constitutional court, he suspended parliament and ruled by decree. New elections in Dec., 1995, gave his allies a majority in parliament but were criticized by the opposition and others as flawed.

plal

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est.

Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast.

The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars.

Despite Kazakhstan's largely arid conditions, its vast steppes accommodate both livestock and grain production.

The Kazakh Hills in the core of the region have important mineral resources.

Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote.

In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army.

Kazakhstan declared its independence from the Soviet Union on Dec.

In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia.

random

Astana is the capital and Almaty (Alma-Ata) is the largest city. Other major cities include Shymkent, Semey, Aqtobe, and Oskemen. Most of the region is desert or has limited and irregular rainfall. The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars. In the 1950s, the Virgin Lands Program under Khrushchev brought hundreds of thousands of Russian, Ukrainian, and German settlers to the area. In addition, there are rich fishing grounds, famous for their caviar-producing sturgeon, in the N Caspian. In the early 1960s parts of republic saw extensive agricultural development as the Virgin Lands Territory. In 1999, as Kazakhstan's economy worsened, the government agreed to sell some of its stake in the vast Tengiz oil field.

sinope

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop.

The population of Kazakhstan consists mainly of Muslim Kazakhs and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars.

The Kazakh Hills in the core of the region have important mineral resources.

In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army.

In 1994, Kazakhstan signed a series of security agreements with the United States, in which the latter would take control of enriched uranium usable for nuclear weapons and aid Kazakhstan in removing extant nuclear weapons, closing missile silos, converting biological-weapons-production centers, and destroying its nuclear test ranges.

In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia.

word

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. Almaty is the site of Kazakhstan Univ. (founded 1934) and the Kazakh Academy of Sciences (founded 1946).

The Baikonur (Bayqongyr) Cosmodrome in central Kazakhstan was the Soviet space-operations center and continues to serve Russian space exploration through an agreement between the two nations. Kazakhstan declared its independence from the Soviet Union on Dec. 16, 1991. In 1994, Kazakhstan signed a series of security agreements with the United States, in which the latter would take control of enriched uranium usable for nuclear weapons and aid Kazakhstan in removing extant nuclear weapons, closing missile silos, converting biological-weapons-production centers, and destroying its nuclear test ranges. In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia. Kazakhstan is a member of the Commonwealth of Independent States.

B.1.2 Long summaries (30%)

copernic

It borders on Siberian Russia in the north, China in the east, Kyrgyzstan, Uzbekistan, Turkmenistan, and the Aral Sea in the south, and the Caspian Sea and European Russia in the west.

Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast.

It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts.

It is largely lowland in the north and west (W Siberian, Caspian, and Turan lowlands), hilly in the center (Kazakh Hills), and mountainous in the south and east (Tian Shan and Altai ranges).

Kazakhstan is a region of inland drainage; the Syr Darya, the Ili, the Chu, and other rivers drain into the Aral Sea and Lake Balkash.

The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars.

There is considerable friction between the now dominant Kazakhs and the formerly favored ethnic Russians, who continue to emigrate in large numbers.

Almaty is the site of Kazakhstan Univ. (founded 1934) and the Kazakh Academy of Sciences (founded 1946).

Despite Kazakhstan's largely arid conditions, its vast steppes accommodate both livestock and grain production.

In the 1950s, the Virgin Lands Program under Khrushchev brought hundreds of thousands of Russian, Ukrainian, and German settlers to the area.

Coal is mined at Qaraghandy and Ekibastuz, and there are major oil fields in the Emba basin (which includes the important Tengiz fields), at the northeast tip of the Caspian Sea, and in the Mangyshlak Peninsula.

Semey was the Soviet center of space-related industries, and the surrounding region was the site of Soviet nuclear testing; radiation pollution is widespread in the area, which experienced a severe economic downturn following the end of nuclear testing in 1991.

Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote.

Nursultan Nazarbayev became the country's first president and soon began a gradual movement toward privatization of the economy.

In Apr., 1995, after the 1994 election results were dismissed as invalid by the constitutional court, he suspended parliament and ruled by decree.

New elections in Dec., 1995, gave his allies a majority in parliament but were criticized by the opposition and others as flawed.

On the basis of referendums held in 1995 and 1996 that were denounced by the opposition, Nazarbayev's term in office was extended to the year 2000 and his powers were increased.

lsa

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. Most of the region is desert or has limited and irregular rainfall. Kazakh, a Turkic language, is the official tongue, but Russian is still widely used.

Almaty is the site of Kazakhstan Univ. (founded 1934) and the Kazakh Academy of Sciences (founded 1946). In the 1950s, the Virgin Lands Program under Khrushchev brought hundreds of thousands of Russian, Ukrainian, and German settlers to the area. The raising of cattle and sheep is also important, and Kazakhstan produces much wool and meat. Kazakhstan also has large deposits of natural gas, iron ore, manganese, chrome, lead, zinc, silver, copper, nickel, titanium, bauxite, and gold. Steel, agricultural and mining machinery, superphosphate fertilizers, phosphorus acids, artificial fibers, synthetic rubber, textiles, and medicines are among the manufactured goods. Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote.

The 19th cent. saw the growth of the Kazakh intelligentsia. A written literature strongly influenced by Russian culture was then developed. In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army. Organized as the Kirghiz Autonomous SSR in 1920, it was renamed the Kazakh Autonomous SSR in 1925 and became a constituent republic in

1936. During the Stalin era, collectivization was instituted and millions of Kazakhs were forced to resettle in the region's south in order to strengthen Russian rule.

In the early 1960s parts of republic saw extensive agricultural development as the Virgin Lands Territory. Nursultan Nazarbayev became the country's first president and soon began a gradual movement toward privatization of the economy.

initial

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. It borders on Siberian Russia in the north, China in the east, Kyrgyzstan, Uzbekistan, Turkmenistan, and the Aral Sea in the south, and the Caspian Sea and European Russia in the west. Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast. It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts. in the east. The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars. Kazakh, a Turkic language, is the official tongue, but Russian is still widely used. Despite Kazakhstan's largely arid conditions, its vast steppes accommodate both livestock and grain production. In the 1950s, the Virgin Lands Program under Khrushchev brought hundreds of thousands of Russian, Ukrainian, and German settlers to the area. The Kazakh Hills in the core of the region have important mineral resources. The country's industries are located along the margins of the country. Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote. The original nomadic Turkic tribes inhabiting the region had a culture that featured the Central Asian epics, ritual songs, and legends. In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army. Kazakhstan declared its independence from the Soviet Union on Dec. Elections in 1994 gave a parliamentary majority to allies of Nazarbayev, but they resisted his reform plans. In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia.

nolsa

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. It borders on Siberian Russia in the north, China in the east, Kyrgyzstan, Uzbekistan, Turkmenistan, and the Aral Sea in the south, and the Caspian Sea and European Russia in the west. Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast. It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts. in the east. It is largely lowland in the north and west (W Siberian, Caspian, and Turan lowlands), hilly in the center (Kazakh Hills), and mountainous in the south and east (Tian Shan and Altai ranges).

In addition, there are rich fishing grounds, famous for their caviar-producing sturgeon, in the N Caspian. Coal is mined at Qaraghandy and Ekibastuz, and there are major oil fields in the Emba basin (which includes the important Tengiz fields), at the northeast tip of the Caspian Sea, and in the Mangyshlak Peninsula. Kashagan, a Caspian field that was being explored in the late 1990s, appears to have great potential. A pipeline was built in the 1990s to connect the nation's oil fields to the Black Sea. The Baikonur (Bayqongyr) Cosmodrome in central Kazakhstan was the Soviet space-operations center and continues to serve Russian space exploration through an agreement between the two nations.

A written literature strongly influenced by Russian culture was then developed. In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army. During the Stalin era, collectivization was instituted and millions of Kazakhs were forced to resettle in the region's south in order to strengthen Russian rule.

These projects were financed by the United States, and many had been completed by late 1999. In Apr., 1995, after the 1994 election results were dismissed as invalid by the constitutional court, he suspended parliament and ruled by decree. New elections in Dec., 1995, gave his allies a majority in parliament but were criticized by the opposition and others as flawed.

plal

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est.

pop.

Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast.

It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts.

The population of Kazakhstan consists mainly of Muslim Kazakhs (more than 45% of the population) and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars.

Kazakh, a Turkic language, is the official tongue, but Russian is still widely used.

Despite Kazakhstan's largely arid conditions, its vast steppes accommodate both livestock and grain production.

In the 1950s, the Virgin Lands Program under Khrushchev brought hundreds of thousands of Russian, Ukrainian, and German settlers to the area.

The Kazakh Hills in the core of the region have important mineral resources.

Coal is mined at Qaraghandy and Ekibastuz, and there are major oil fields in the Emba basin (which includes the important Tengiz fields), at the northeast tip of the Caspian Sea, and in the Mangyshlak Peninsula.

The country's industries are located along the margins of the country.

Under the constitution of 1995, Kazakhstan is headed by a strong executive president, who is elected by popular vote.

There is a bicameral parliament, most of whose members are elected, but its powers are limited.

The original nomadic Turkic tribes inhabiting the region had a culture that featured the Central Asian epics, ritual songs, and legends.

In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army.

Kazakhstan declared its independence from the Soviet Union on Dec.

Elections in 1994 gave a parliamentary majority to allies of Nazarbayev, but they resisted his [Caspian] reform plans.

In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia.

random

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. It extends nearly 2,000 mi (3,200 km) from the lower Volga and the Caspian Sea in the west to the Altai Mts. in the east. Kazakhstan is a region of inland drainage; the Syr Darya, the Ili, the Chu, and other rivers drain into the Aral Sea and Lake Balkash. Kazakh, a Turkic language, is the official tongue, but Russian is still widely used. There is considerable friction between the now dominant Kazakhs and the formerly favored ethnic Russians, who continue to emigrate in large numbers. Wheat, cotton, sugar beets, and tobacco are the main crops. In addition, there are rich fishing grounds, famous for their caviar-producing sturgeon, in the N Caspian. A pipeline was built in the 1990s to connect the nation's oil fields to the Black Sea. Kazakhstan also has large deposits of natural gas, iron ore, manganese, chrome, lead, zinc, silver, copper, nickel, titanium, bauxite, and gold. Steel, agricultural and mining machinery, superphosphate fertilizers, phosphorus acids, artificial fibers, synthetic rubber, textiles, and medicines are among the manufactured goods. The main trading partners are Russia, Ukraine, and Uzbekistan. The original nomadic Turkic tribes inhabiting the region had a culture that featured the Central Asian epics, ritual songs, and legends. The 19th cent. saw the growth of the Kazakh intelligentsia. Elections in 1994 gave a parliamentary majority to allies of Nazarbayev, but they resisted his reform plans. On the basis of referendums held in 1995 and 1996 that were denounced by the opposition, Nazarbayev's term in office was extended to the year 2000 and his powers were increased. In 1999, as Kazakhstan's economy worsened, the government agreed to sell some of its stake in the vast Tengiz oil field.

sinope

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop.

Kazakhstan consists of a vast flatland, bordered by a high mountain belt in the southeast. Kazakhstan is a region of inland drainage; the Syr Darya, the Ili, the Chu, and other rivers drain into the Aral Sea and Lake Balkash.

The population of Kazakhstan consists mainly of Muslim Kazakhs and Russians (some 35%, many of whom belong to the Russian Orthodox Church); there are smaller minorities of Ukrainians, Germans, Uzbeks, and Tatars.

The Kazakh Hills in the core of the region have important mineral resources. Kazakhstan also has large deposits of natural gas, iron ore, manganese, chrome, lead, zinc, silver, copper, nickel, titanium, bauxite, and gold.

The country's industries are located along the margins of the country. The Baikonur Cosmodrome in central Kazakhstan was the Soviet space-operations center and continues to serve Russian space exploration through an agreement between the two nations. The main trading partners are Russia, Ukraine, and Uzbekistan.

The original nomadic Turkic tribes inhabiting the region had a culture that featured the Central Asian epics, ritual songs, and legends.

In 1916 the Kazakhs rebelled against Russian domination and were in the process of establishing a Western-style state at the time of the 1917 Bolshevik Revolution, but by 1920 the region was under the control of the Red Army.

Kazakhstan declared its independence from the Soviet Union on Dec. In 1994, Kazakhstan signed a series of security agreements with the United States, in which the latter would take control of enriched uranium usable for nuclear weapons and aid Kazakhstan in removing extant nuclear weapons, closing missile silos, converting biological-weapons-production centers, and destroying its nuclear test ranges.

In an election rescheduled to Jan., 1999, Nazarbayev was reelected after disqualifying the major opposition candidate.

In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia. Kazakhstan is a member of the Commonwealth of Independent States.

word

Kazakhstan, or Kazakstan, officially Republic of Kazakhstan, republic (1995 est. pop. 17,377,000), c.1,050,000 sq mi (2,719,500 sq km), central Asia. It borders on Siberian Russia in the north, China in the east, Kyrgyzstan, Uzbekistan, Turkmenistan, and the Aral Sea in the south, and the Caspian Sea and European Russia in the west. Kazakhstan is a region of inland drainage; the Syr Darya, the Ili, the Chu, and other rivers drain into the Aral Sea and Lake Balkash. Kazakh, a Turkic language, is the official tongue, but Russian is still widely used. Almaty is the site of Kazakhstan Univ. (founded 1934) and the Kazakh Academy of Sciences (founded 1946).

The Kazakh Hills in the core of the region have important mineral resources. Kazakhstan also has large deposits of natural gas, iron ore, manganese, chrome, lead, zinc, silver, copper, nickel, titanium, bauxite, and gold. The country's industries are located along the margins of the country. The Baikonur (Bayqongyr) Cosmodrome in central Kazakhstan was the Soviet space-operations center and continues to serve Russian space exploration through an agreement between the two nations. The 19th cent. A written literature strongly influenced by Russian culture was then developed.

Kazakhstan declared its independence from the Soviet Union on Dec. 16, 1991. In 1994, Kazakhstan signed a series of security agreements with the United States, in which the latter would take control of enriched uranium usable for nuclear weapons and aid Kazakhstan in removing extant nuclear weapons, closing missile silos, converting biological-weapons-production centers, and destroying its nuclear test ranges. In an election rescheduled to Jan., 1999, Nazarbayev was reelected after disqualifying the major opposition candidate.

In 1996, Kazakhstan, along with Kyrgyzstan and Belarus, signed an economic cooperation pact with Russia. In 1999, as Kazakhstan's economy worsened, the government agreed to sell some of its stake in the vast Tengiz oil field. Kazakhstan is a member of the Commonwealth of Independent States.

B.2 pres92

B.2.1 Short summaries (5%)

copernic

For voters unhappy with President Bush's passive approach to the economy, the six main contenders for the Democratic nomination offer a smorgasbord of proposals.

There are tax cuts for the middle class; tax incentives for investment and savings; tax breaks for families with children; and spending on health, education, roads, high-speed transportation, high-tech telecommunications and much more.

Among the Democratic candidates, he expresses the least concern about the budget deficit.

Mr. Tsongas, meanwhile, favors "strategic" government investments in technology and targeted tax cuts for investment—including a capital-gains cut.

lsa

Over the past decade, the party's presidential nominees have felt compelled to battle against the giant federal budget deficits created under President Reagan. Neither was able to sell that sour medicine to a public averse to pain. Focusing the candidates' discussion of economic policy are three separate questions: How can the government get the economy out of recession in the near term? Mr Clinton's version calls for a cut of about $350 a year for middle-class taxpayers.

initial

Over the past decade, the party's presidential nominees have felt compelled to battle against the giant federal budget deficits created under President Reagan. This year will be different. As a result, the Democratic candidates are no longer paralyzed by the deficit issue. "They are all much more willing in 1992 to talk about government responsibility for the economy than either Dukakis or Mondale," says Jeff Faux, president of the liberal Economic Policy Institute.

nolsa

Over the past decade, the party's presidential nominees have felt compelled to battle against the giant federal budget deficits created under President Reagan. And the collapse of the Soviet empire has created promises of a pot of gold, perhaps as much as $150 billion a year, at the end of the defense-cutting rainbow.

Focusing the candidates' discussion of economic policy are three separate questions: How can the government get the economy out of recession in the near term? To address long-term problems, he favors an investment tax credit, a targeted capital-gains tax cut for investments in new businesses, and more government spending on civilian research.

plal

Over the past decade, the party's presidential nominees have felt compelled to battle against the giant federal budget deficits created under President Reagan.

This year will be different.

As a result, the Democratic candidates are no longer paralyzed by the deficit issue.

"They are all much more willing in 1992 to talk about government responsibility for the economy than either Dukakis or Mondale," says Jeff Faux, president of the liberal Economic Policy Institute.

Focusing the candidates' discussion of economic policy are three separate questions: How can the government get the economy out of recession in the near term?

random

Just how to get the money out of overhead, or which programs are low-priority, is unclear. He would encourage them through tax incentives and through "strategic" investments in technology. But instead of cutting the middle class's taxes, he seeks a temporary investment tax credit. Many of the spending plans Mr Harkin envisions on a grand scale are echoed on a smaller scale in speeches by Mr Kerrey and Mr Clinton.

sinope

So far, the most popular answer is the middle-class tax cut; it is embraced, in some form, by four of the six candidates: Govs.

To pay for all these, he'd propose a 3% across-the-board cut in all government administrative expenses—a proposal that's popular among politicians seeking office, but seldom practical once they get there.

The other candidate who has put tax cuts at the heart of his campaign is Mr.

Tsongas also calls on the Fed to lower interest rates, saying that recent cuts " haven't worked since rates are still too high in relation to inflation."

word

Mr. Tsongas, meanwhile, favors "strategic" government investments in technology and targeted tax cuts for investment—including a capital-gains cut.

To address long-term problems, he favors an investment tax credit, a targeted capital-gains tax cut for investments in new businesses, and more government spending on civilian research.

Mr. Wilder combines his tax-cut proposal with broad but general calls for government spending reductions. Mr. Kerrey's middle-class tax cut, like Mr. Clinton's, would be paid for by higher taxes on the wealthy.

B.3 civilwar

B.3.1 Short summaries (5%)

copernic

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth.

Most simply stated, the constitutional issue was one between a king who claimed to rule by divine right and a Parliament that professed itself to have rights and privileges independent of the crown and that ultimately, by its actions, claimed real sovereignty.

Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber.

This distrust was given sharp focus by the outbreak (Oct., 1641) of a rebellion against English rule in Ireland; an army was needed to suppress the rebellion, but the parliamentarians feared that the king might use it against them.

lsa

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. That name, however, overemphasizes the religious element at the expense of the constitutional issues and the underlying social and economic factors. The Parliament of 1626 went further and impeached the king's favorite, George Villiers, 1st duke of Buckingham. Those imprisoned by the Star Chamber were freed.

initial

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy. Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans. James I was not long in gaining a personal unpopularity that helped to strengthen Parliament's hand.

nolsa

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. The Parliament of 1625 granted him the right to collect tonnage and poundage (customs duties) only for a year and not, as was customary, for his entire reign. Those imprisoned by the Star Chamber were freed. The radicalism of these demands split the parliamentary party and drove many of the moderates to the royalist side.

plal

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth.

The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy.

Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans.

James I was not long in gaining a personal unpopularity that helped to strengthen Parliament's hand.

random

The king dismissed Coke from the bench in 1616, but the Parliament of 1621 impeached Bacon. The ominous peace was broken by troubles in Scotland, where efforts to enforce Anglican episcopal policy led to the violent opposition of the Covenanters and to war in 1639 (see Bishops' Wars) and compelled Charles to seek the financial aid of Parliament. Ship money and tonnage and poundage without parliamentary authorization were abolished. The army resisted Parliament's proposal to disband it by capturing the king from the parliamentary party and marching on London.

sinope

Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans.

In June, 1642, Parliament sent to the king a statement reiterating the demands of the Grand Remonstrance, but since the proposals amounted to a complete surrender of sovereignty by the crown to Parliament, the king did not even consider them as a basis for discussion.

The king was delivered by the Scots into the hands of Parliament, but the Presbyterian rule in that body had thoroughly alienated the army.

Parliament again tried to reach some agreement with the king, but the army, now completely under Cromwell's domination, disposed of its enemies in Parliament by Pride's Purge.

word

The Parliament that met in 1604 soon clashed with the king on questions of finance and supply. The king dismissed Coke from the bench in 1616, but the Parliament of 1621 impeached Bacon. The Parliament of 1626 went further and impeached the king's favorite, George Villiers, 1st duke of Buckingham. Charles dissolved it in anger.

Charles offered to abandon ship money exactions, but the opposition wished to discuss more fundamental issues, and the king dissolved the Parliament in just three weeks.

In the lull that followed, both Parliament and the king sought to secure fortresses, arsenals, and popular support. Charles then established himself at Oxford. The army resisted Parliament's proposal to disband it by capturing the king from the parliamentary party and marching on London.

B.3.2 Long summaries (15%)

copernic

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth.

The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy.

Most simply stated, the constitutional issue was one between a king who claimed to rule by divine right and a Parliament that professed itself to have rights and privileges independent of the crown and that ultimately, by its actions, claimed real sovereignty.

Their ambition to do away with financial and commercial restrictions and their desire to have a say in such matters as religious and foreign policies had been severely restrained by the Tudors, but on the accession (1603) of a Scottish king to the English throne the popular party began to organize its strength.

James was forced to temporize because of his urgent need of money, but the dissolution of the Parliament in 1610 left feelings of bitterness on both sides.

James had little understanding of the popular unrest and aroused deeper opposition by his continued collection of impositions and benevolences, his dependence on favorites, and his scheme of a Spanish marriage for his son Charles.

Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber.

The religious issue also came up, and Commons resisted the king's order to adjourn by forcing the speaker to remain in his chair while Eliot presented resolutions against "popery" and unauthorized taxation.

This distrust was given sharp focus by the outbreak (Oct., 1641) of a rebellion against English rule in Ireland; an army was needed to suppress the rebellion, but the parliamentarians feared that the king might use it against them.

The indecisive engagements of 1643 were remarkable mainly for the emergence of Oliver Cromwell, an inconspicuous member of the Long Parliament, to military prominence with his own regiment of "godly" men, soon to become famous as the Ironsides.

Futile negotiations for peace had been conducted at Oxford early in 1643, and in Sept., 1643, Parliament took a decisive step by securing the alliance of the Presbyterian Scots in accepting the Solemn League and Covenant.

His campaign to relieve the besieged York led to the battle of Marston Moor (July 2, 1644), in which Cromwell and Leslie inflicted a crushing defeat on the royalists.

Charles managed to cut off Essex in the southwest but shortly thereafter met parliamentary troops from the north in an indecisive engagement at Newbury.

lsa

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy. That name, however, overemphasizes the religious element at the expense of the constitutional issues and the underlying social and economic factors. The Parliament that met in 1604 soon clashed with the king on questions of finance and supply. The Parliament of 1626 went further and impeached the king's favorite, George Villiers, 1st duke of Buckingham. Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber. Those imprisoned by the Star Chamber were freed. Strafford was impeached, then attainted and executed (1641) for treason; Laud was impeached and imprisoned. Despite the king's compliance to the will of the opposition thus far, he was not trusted by the parliamentary party. This encouraged Charles to assert himself, and in Jan., 1642, he attempted to arrest in person Pym and four other leaders of the opposition in Commons. Charles managed to cut off Essex in the southwest but shortly thereafter met parliamentary troops from the north in an indecisive engagement at Newbury. The legislative remnant known as the Rump Parliament erected a high court of justice, which tried the king for treason and found him guilty. Charles was beheaded on Jan. 30, 1649, and the republic known as the Commonwealth was set up, governed by the Rump Parliament (without the House of Lords) and by an executive council of state.

initial

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy. Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans. James I was not long in gaining a personal unpopularity that helped to strengthen Parliament's hand. A new Parliament met in 1614, and the Commons engaged in quarrels not only with the king but also with the House of Lords. Meanwhile a legal battle was being waged in the courts, with Sir Francis Bacon zealously upholding the royal prerogative and Sir Edward Coke defending the supremacy of common law. Charles I, married to a French Roman Catholic princess, Henrietta Maria, proved more intractable and even less acceptable to the Puritan taste than his father, and Parliament became even more uncompromising in the new reign. Failing to raise money without Parliament, he was forced to call a new one in 1628. In the succeeding 11 years Charles attempted to rule without a Parliament, resorting to such expedients as ship money (a tax levied originally on seaports but extended by Charles to the entire country) to raise revenue. The ominous peace was broken by troubles in Scotland, where efforts to enforce Anglican episcopal policy led to the violent opposition of the Covenanters and to war in 1639 (see Bishops' Wars) and compelled Charles to seek the financial aid of Parliament. The disasters of the second Scottish war compelled a virtual surrender by the king to the opposition, and the Long Parliament was summoned (Nov. Despite the king's compliance to the will of the opposition thus far, he was not trusted by the parliamentary party. In the lull that followed, both Parliament and the king sought to secure fortresses, arsenals, and popular support.

nolsa

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth. Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans. The 16th cent. had seen a decline in the influence of the nobility and a striking rise in the numbers, wealth, and influence of the gentry and merchants, the beneficiaries of a tremendous expansion of markets and trade in Tudor times. James had little understanding of the popular unrest and aroused deeper opposition by his continued collection of impositions and benevolences, his dependence on favorites, and his scheme of a Spanish marriage for his son Charles. The Parliament of 1625 granted him the right to collect tonnage and poundage (customs duties) only for a year and not, as was customary, for his entire reign. Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber. Those imprisoned by the Star Chamber were freed.

Ship money and tonnage and poundage without parliamentary authorization were abolished. The radicalism of these demands split the parliamentary party and drove many of the moderates to the royalist side. Armed forces (including many peers from the House of Lords and a sizable minority of Commons) gathered about him in the north. A Scottish army, under Alexander Leslie, 1st earl of Leven, advanced into Yorkshire early in 1644 and gave aid to the parliamentary army in the north. Unable to join Montrose (who was defeated by Leslie in Scotland) and thwarted in his attempts to secure aid from Ireland or the Continent, the king was unable to halt the steady losses of his party and finally was compelled to surrender himself to the Scots, who made him reassuring but vague promises. Charles I's son Charles II was recognized as king in parts of Ireland and in Scotland but was forced to flee to the Continent after his defeat at Worcester (1651).

plal

English civil war, 1642–48, the conflict between King Charles I of England and a large body of his subjects, generally called the "parliamentarians," that culminated in the defeat and execution of the king and the establishment of a republican commonwealth.

The struggle has also been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy.

Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans.

James I was not long in gaining a personal unpopularity that helped to strengthen Parliament's hand.

A new Parliament met in 1614, and the Commons engaged in quarrels not only with the king but also with the House of Lords.

Charles I, married to a French Roman Catholic princess, Henrietta Maria, proved more intractable and even less acceptable to the Puritan taste than his father, and Parliament became even more uncompromising in the new reign.

In the succeeding 11 years Charles attempted to rule without a Parliament, resorting to such expedients as ship money (a tax levied originally on seaports but extended by Charles to the entire country) to raise revenue.

The ominous peace was broken by troubles in Scotland, where efforts to enforce Anglican episcopal policy led to the violent opposition of the Covenanters and to war in 1639 (see Bishops' Wars) and compelled Charles to seek the financial aid of Parliament.

The disasters of the second Scottish war compelled a virtual surrender by the king to the opposition, and the Long Parliament was summoned (Nov.

In the lull that followed, both Parliament and the king sought to secure fortresses, arsenals, and popular support.

The followers of king and Parliament did not represent two absolutely distinct social groups, as the popular conception of the royalist Cavaliers and the parliamentary Roundheads would indicate.

The king was delivered (1647) by the Scots into the hands of Parliament, but the Presbyterian rule in that body had thoroughly alienated the army.

random

Meanwhile a legal battle was being waged in the courts, with Sir Francis Bacon zealously upholding the royal prerogative and Sir Edward Coke defending the supremacy of common law. The last Parliament (1624) of the reign accompanied its grant of money with specific directions for its use. The Parliament of 1625 granted him the right to collect tonnage and poundage (customs duties) only for a year and not, as was customary, for his entire reign. The Parliament of 1626 went further and impeached the king's favorite, George Villiers, 1st duke of Buckingham. The resulting Short Parliament (1640) once more met the king's request for supply by a demand for redress of grievance. Charles offered to abandon ship money exactions, but the opposition wished to discuss more fundamental issues, and the king dissolved the Parliament in just three weeks. The parliamentarians quickly enacted a series of measures designed to sweep away what they regarded as the encroachments of despotic monarchy. This encouraged Charles to assert himself, and in Jan., 1642, he attempted to arrest in person Pym and four other leaders of the opposition in Commons. Thomas Fairfax (later 3d Baron Fairfax of Cameron) became the commander in chief. After further futile peace negotiations at Uxbridge, Charles, hoping to join the forces under James Graham, marquess of Montrose, moved north and stormed Leicester. This battle cost the king a large part of his army and rendered the royalist cause hopeless. In the spring of 1648, the second civil war began. Charles was beheaded on Jan. 30, 1649, and the republic known as the Commonwealth was set up, governed by the Rump Parliament (without the House of Lords) and by an executive council of state.

sinope

The struggle has been called the Puritan Revolution because the religious complexion of the king's opponents was prevailingly Puritan, and because the defeat of the king was accompanied by the abolition of episcopacy.

Parliament in this period did not represent the full body of the English people; it was composed of and represented the nobility, country gentry, and merchants and artisans.

A new Parliament met in 1614, and the Commons engaged in quarrels not only with the king but also with the House of Lords.

The religious issue also came up, and Commons resisted the king's order to adjourn by forcing the speaker to remain in his chair while Eliot presented resolutions against " popery" and unauthorized taxation.

In the succeeding 11 years Charles attempted to rule without a Parliament, resorting to such expedients as ship money to raise revenue.

The ominous peace was broken by troubles in Scotland, where efforts to enforce Anglican episcopal policy led to the violent opposition of the Covenanters and to war in 1639 and compelled Charles to seek the financial aid of Parliament.

The disasters of the second Scottish war compelled a virtual surrender by the king to the opposition, and the Long Parliament was summoned.

In June, 1642, Parliament sent to the king a statement reiterating the demands of the Grand Remonstrance, but since the proposals amounted to a complete surrender of sovereignty by the crown to Parliament, the king did not even consider them as a basis for discussion.

The followers of king and Parliament did not represent two absolutely distinct social groups, as the popular conception of the royalist Cavaliers and the parliamentary Roundheads would indicate.

The king was delivered by the Scots into the hands of Parliament, but the Presbyterian rule in that body had thoroughly alienated the army.

Parliament again tried to reach some agreement with the king, but the army, now completely under Cromwell's domination, disposed of its enemies in Parliament by Pride's Purge.

word

The Parliament that met in 1604 soon clashed with the king on questions of finance and supply. The king dismissed Coke from the bench in 1616, but the Parliament of 1621 impeached Bacon. The Parliament of 1626 went further and impeached the king's favorite, George Villiers, 1st duke of Buckingham. Charles dissolved it in anger.

The new Parliament drew up the Petition of Right, and Charles accepted it in order to get his subsidy. Parliament in 1629 vigorously protested Charles's collection of tonnage and poundage and the prosecution of his opponents in the Star Chamber. The resulting Short Parliament (1640) once more met the king's request for supply by a demand for redress of grievance. Charles offered to abandon ship money exactions, but the opposition wished to discuss more fundamental issues, and the king dissolved the Parliament in just three weeks.

The disasters of the second Scottish war compelled a virtual surrender by the king to the opposition, and the Long Parliament was summoned (Nov., 1640). Led by John Pym, Parliament adopted the Grand Remonstrance, reciting the evils of Charles's reign and demanding church reform and parliamentary control over the army and over the appointment of royal ministers. In the lull that followed, both Parliament and the king sought to secure fortresses, arsenals, and popular support. On Aug. 22, 1642, Charles raised his standard at Nottingham.

Charles then established himself at Oxford. To stem the rising dissension among parliamentary leaders, Cromwell sponsored in Parliament the Self-Denying Ordinance, by which all members of Parliament were compelled to resign their commands, and the parliamentary army was reorganized (1644–45) into the New Model Army. The king was delivered (1647) by the Scots into the hands of Parliament, but the Presbyterian rule in that body had thoroughly alienated the army. The army resisted Parliament's proposal to disband it by capturing the king from the parliamentary party and marching on London. Parliament again tried to reach some agreement with the king, but the army, now completely under Cromwell's domination, disposed of its enemies in Parliament by Pride's Purge (Dec., 1648; see under Pride, Thomas).

Appendix C

Judges' ratings

Table C.1: Raw coherence ratings

Source	Length	Summarizer	Judge																	
			A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
kazakh	short	nolsa	3	2	2	2	5	3	1	2	4	4	1	3	3	3	3	3	3	3
kazakh	short	random	2	4	4	4	5	2	3	5	5	2	4	4	4	3	3	3	2	4
kazakh	short	sinope	4	4	5	4	2	3	4	4	3	2	4	3	4	4	4	4	2	4
kazakh	short	copern	3	4	3	4	3	4	2	3	5	2	4	4	4	3	3	4	4	5
kazakh	short	word	2	5	3	3	4	3	4	2	3	4	3	4	4	3	3	2	2	3
kazakh	short	lsa	3	3	4	3	2	2	2	5	5	4	4	2	2	3	3	2	3	3
kazakh	short	init	4	5	4	5	5	3	4	3	3	4	3	2	4	3	4	3	2	4
kazakh	short	plal	4	5	5	4	3	2	3	5	4	3	3	4	4	3	3	4	4	2
kazakh	long	lsa	3	3	3	4	2	4	2	3	4	2	2	2	4	5	4	4	2	5
kazakh	long	sinope	4	4	5	3	1	2	3	1	4	4	3	4	4	3	5	5	4	4
kazakh	long	init	2	4	2	4	5	4	3	4	4	3	4	2	5	3	5	4	2	3
kazakh	long	copern	2	5	4	4	3	2	1	5	3	3	5	4	3	4	4	4	2	3
kazakh	long	plal	3	5	5	2	4	1	1	3	4	2	3	1	3	3	3	3	3	5
kazakh	long	random	3	3	3	3	4	4	1	5	4	4	3	4	4	4	3	3	2	5
kazakh	long	word	4	4	4	4	2	3	4	1	5	4	4	3	1	4	3	3	2	5
kazakh	long	nolsa	3	4	3	2	1	2	1	2	5	3	3	2	1	4	2	2	4	5
pres92	short	lsa	3	5	4	1	4	2	5	3	3	2	4	1	3	4	2	3	2	5
pres92	short	copern	3	3	2	3	3	3	2	5	4	2	3	2	1	4	3	2	2	4
pres92	short	plal	3	2	2	4	1	3	1	4	4	1	4	2	2	4	2	3	2	2
pres92	short	nolsa	2	3	2	2	4	3	1	2	4	4	1	2	3	5	2	3	5	4
pres92	short	random	3	2	1	3	2	2	5	5	2	4	3	1	1	3	1	3	1	1
pres92	short	sinope	2	3	1	1	2	1	2	3	1	2	2	1	1	3	1	2	2	3
pres92	short	word	3	3	2	3	3	3	4	5	4	3	2	3	2	4	2	2	1	3
pres92	short	init	3	5	4	4	1	4	5	4	5	4	5	4	3	5	4	2	3	5
pres92	long	nolsa	3	3	3	3	4	4	2	5	5	2	5	3	3	4	4	2	2	4
pres92	long	word	3	3	4	4	3	2	3	3	5	3	4	4	2	3	3	3	4	2

Table C.1: Raw coherence ratings (continued)

Source	Length	Summarizer	Judge																	
			A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
pres92	long	lsa	2	2	4	1	3	2	2	2	5	1	4	3	2	4	3	2	3	4
pres92	long	init	4	4	4	4	4	3	4	5	5	2	5	3	4	4	4	3	4	5
pres92	long	sinope	2	1	1	1	2	1	2	1	1	1	1	2	1	2	1	1	2	1
pres92	long	random	4	1	3	2	3	2	2	2	3	4	5	3	3	4	3	2	2	5
pres92	long	plal	4	3	4	4	4	3	3	5	5	4	5	2	4	3	5	3	4	5
pres92	long	copern	3	3	5	1	2	3	2	5	4	3	4	2	1	2	3	2	4	4
civil	short	random	2	2	4	1	3	2	1	4	1	1	3	2	2	4	1	1	1	5
civil	short	init	4	2	4	4	4	4	1	5	4	3	5	4	4	4	5	3	4	5
civil	short	copern	4	3	5	3	3	2	1	5	4	4	4	2	2	4	3	3	4	4
civil	short	lsa	3	1	2	2	1	2	1	4	2	3	1	2	2	4	3	2	2	4
civil	short	word	2	4	2	2	1	2	1	4	3	3	3	2	4	3	2	3	1	4
civil	short	sinope	1	3	4	2	1	3	1	4	4	1	1	2	3	3	1	2	2	3
civil	short	nolsa	2	1	4	2	1	3	1	3	2	1	2	4	2	4	2	2	1	4
civil	short	plal	4	3	5	3	3	3	1	4	4	3	5	2	4	3	4	3	4	4
civil	long	lsa	2	1	5	2	3	3	1	5	2	3	3	3	4	4	2	3	4	5
civil	long	word	3	1	3	1	3	2	1	5	2	1	4	4	3	3	4	3	3	3
civil	long	random	2	3	2	3	2	2	2	5	2	2	2	2	3	3	1	3	2	3
civil	long	plal	3	4	5	4	3	3	1	4	3	4	5	2	2	2	5	4	4	4
civil	long	nolsa	2	3	5	3	4	2	1	4	5	2	3	3	3	4	4	3	4	4
civil	long	sinope	3	2	4	3	3	4	1	3	3	1	5	2	2	4	4	3	3	4
civil	long	copern	3	3	4	3	3	3	1	5	3	3	3	2	4	3	4	4	3	5
civil	long	init	4	2	3	4	2	4	2	3	4	3	5	2	5	4	4	4	2	5

Table C.2: Raw comprehensiveness ratings

Source	Length	Summarizer	Judge A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
kazakh	short	nolsa	2	4	4	4	3	4	1	4	4	4	2	4	4	3	3	4	2	3
kazakh	short	random	2	1	1	1	4	2	2	4	1	2	1	2	2	2	3	3	1	3
kazakh	short	sinope	3	3	5	3	3	2	3	5	4	3	3	2	3	3	4	3	2	4
kazakh	short	copern	4	3	3	3	3	4	3	3	4	3	3	3	4	4	3	4	5	4
kazakh	short	word	2	1	2	1	3	2	3	2	1	3	3	2	2	2	3	2	3	3
kazakh	short	lsa	2	3	2	4	4	3	2	4	4	4	2	2	3	2	4	2	3	4
kazakh	short	init	3	4	5	1	3	3	3	3	4	4	4	2	1	4	4	3	3	5
kazakh	short	plal	4	4	5	5	4	4	5	5	5	4	5	3	5	5	4	3	5	4
kazakh	long	lsa	3	2	3	2	4	2	3	4	3	2	3	3	3	3	4	3	3	4
kazakh	long	sinope	4	3	5	4	4	4	4	4	5	4	4	4	5	5	5	4	4	4
kazakh	long	init	3	4	2	2	5	4	4	4	5	5	3	2	4	5	5	5	4	3
kazakh	long	copern	2	3	4	4	4	4	5	5	5	3	2	4	3	5	5	4	5	4
kazakh	long	plal	3	5	5	2	5	3	5	4	5	4	2	2	5	4	4	4	4	4
kazakh	long	random	2	4	3	2	2	4	4	5	4	4	3	4	3	4	3	4	2	4
kazakh	long	word	3	3	4	3	3	3	2	3	4	2	4	3	4	3	3	3	3	4
kazakh	long	nolsa	2	3	3	4	3	3	1	3	3	3	4	3	3	4	2	3	4	3
pres92	short	lsa	1	3	2	2	4	2	2	4	3	2	3	1	1	3	3	2	3	2
pres92	short	copern	4	3	4	3	4	2	3	4	4	2	4	3	2	4	4	3	4	3
pres92	short	plal	3	2	3	1	2	3	2	3	4	2	4	3	2	3	3	2	2	2
pres92	short	nolsa	2	2	1	2	3	2	1	3	4	2	2	2	2	4	2	2	4	3
pres92	short	random	2	2	1	1	2	2	3	3	2	1	1	1	1	2	1	1	1	2
pres92	short	sinope	2	1	2	3	1	2	2	1	2	2	1	2	1	3	1	1	1	3
pres92	short	word	2	3	1	3	2	3	3	4	3	2	2	1	3	2	2	1	1	4
pres92	short	init	3	3	3	1	2	2	1	3	5	2	3	3	1	2	4	2	2	2
pres92	long	nolsa	2	2	4	2	2	1	3	1	3	2	3	2	1	2	4	3	3	5
pres92	long	word	4	3	4	5	3	3	4	4	4	3	3	4	4	4	3	2	2	4

Table C.2: Raw comprehensiveness ratings (continued)

Source	**Length**	**Summarizer**	**Judge**																	
			A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**
pres92	long	lsa	2	2	3	2	2	2	3	2	4	2	4	2	3	3	3	3	2	4
pres92	long	init	3	3	4	4	4	3	5	5	5	3	4	3	3	4	4	3	2	4
pres92	long	sinope	3	3	3	4	2	2	4	1	3	2	3	2	2	3	1	1	1	4
pres92	long	random	4	2	4	4	4	2	3	3	3	3	3	2	2	2	3	2	2	4
pres92	long	plal	3	3	5	4	5	4	4	4	5	4	5	4	5	4	5	3	4	5
pres92	long	copern	3	3	5	2	3	2	3	5	5	3	4	2	3	3	3	2	3	5
civil	short	random	1	1	2	3	1	2	2	3	2	2	1	3	1	2	1	1	1	2
civil	short	init	3	2	4	4	1	3	2	2	4	3	3	4	1	2	5	2	2	4
civil	short	copern	4	3	5	3	4	3	1	5	3	5	2	2	3	3	3	2	3	4
civil	short	lsa	2	2	2	3	2	2	2	2	2	2	2	2	2	2	3	2	2	3
civil	short	word	1	1	3	3	2	3	3	4	3	1	1	3	2	4	2	2	2	3
civil	short	sinope	1	2	2	1	3	3	2	2	4	1	3	2	2	2	1	2	1	3
civil	short	nolsa	2	2	4	1	2	2	2	3	2	1	2	4	3	3	2	3	3	3
civil	short	plal	3	2	5	3	2	3	2	1	4	3	3	2	1	3	4	3	2	3
civil	long	lsa	3	2	5	4	2	4	2	4	3	3	4	3	4	3	2	4	3	4
civil	long	word	2	2	2	4	2	3	2	4	4	2	2	3	5	2	4	2	4	4
civil	long	random	2	3	3	4	1	3	3	5	3	2	3	3	3	3	1	3	3	5
civil	long	plal	3	4	5	3	2	4	4	5	5	4	4	3	4	4	5	3	2	5
civil	long	nolsa	3	3	5	2	3	3	3	5	5	3	4	3	3	2	4	2	4	5
civil	long	sinope	3	2	2	2	2	4	4	4	4	2	4	2	3	2	3	3	4	4
civil	long	copern	4	3	4	2	2	3	4	5	4	3	3	2	3	4	3	3	4	4
civil	long	init	3	2	3	2	2	3	2	4	3	3	4	2	2	3	3	2	2	4

Table C.3: Raw overall quality ratings

Source	Length	Summarizer	Judge																	
			A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
kazakh	short	nolsa	3	3	4	2	4	4	1	2	5	4	2	3	3	3	4	4	3	3
kazakh	short	random	2	2	3	2	4	2	1	4	3	2	3	4	3	2	4	3	2	3
kazakh	short	sinope	4	3	5	3	3	2	2	5	4	3	3	3	4	3	4	3	2	4
kazakh	short	copern	4	4	4	3	2	4	3	2	5	2	3	4	4	4	4	4	4	5
kazakh	short	word	2	2	2	1	4	2	2	1	4	3	4	3	3	2	4	2	3	3
kazakh	short	lsa	2	2	3	3	2	2	1	5	5	4	3	1	2	3	4	2	2	4
kazakh	short	init	3	4	4	2	4	3	4	3	4	4	3	2	3	3	5	3	4	4
kazakh	short	plal	4	5	5	4	3	3	5	5	5	3	4	3	4	5	4	3	4	2
kazakh	long	lsa	3	4	4	3	3	3	3	4	3	2	2	3	3	3	4	4	3	4
kazakh	long	sinope	4	3	5	4	2	4	4	2	5	4	4	3	4	5	5	5	4	4
kazakh	long	init	3	4	2	2	5	3	4	3	5	4	3	1	5	4	5	5	3	3
kazakh	long	copern	2	3	4	4	2	4	4	5	4	3	3	5	2	5	4	4	4	3
kazakh	long	plal	3	5	3	3	4	2	4	4	5	4	2	2	3	4	4	3	4	4
kazakh	long	random	3	4	3	2	3	4	3	5	3	4	2	4	3	4	3	4	2	5
kazakh	long	word	3	3	4	4	2	3	3	1	4	3	4	3	2	3	4	3	3	5
kazakh	long	nolsa	2	2	3	4	1	2	1	1	3	3	3	2	1	4	2	3	4	4
pres92	short	lsa	2	2	3	1	4	2	4	4	3	2	3	1	2	3	3	2	3	3
pres92	short	copern	4	4	4	2	3	3	2	4	4	2	3	3	1	4	4	2	3	3
pres92	short	plal	3	2	2	1	1	2	1	1	5	2	4	3	2	4	3	2	2	2
pres92	short	nolsa	2	1	2	2	3	2	1	1	4	3	1	2	2	5	2	2	4	3
pres92	short	random	2	1	1	1	2	1	2	3	1	1	1	1	1	2	1	1	1	1
pres92	short	sinope	2	1	1	2	1	1	2	2	1	1	1	1	1	3	1	1	1	4
pres92	short	word	2	2	2	2	2	3	3	5	4	2	2	2	2	3	2	1	1	4
pres92	short	init	3	4	4	1	1	3	1	4	5	2	4	3	2	3	4	2	2	4
pres92	long	nolsa	2	2	4	2	3	3	2	4	3	2	4	2	2	3	5	3	2	4
pres92	long	word	3	2	4	4	3	2	3	4	4	3	3	4	3	4	3	3	3	3

Table C.3: Raw overall quality ratings (continued)

Source	Length	Summarizer	Judge																	
			A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**	**N**	**O**	**P**	**Q**	**R**
pres92	long	lsa	2	2	4	2	2	2	2	2	4	2	4	2	2	3	4	3	2	4
pres92	long	init	4	3	4	4	4	3	5	5	5	3	5	3	3	4	4	3	3	4
pres92	long	sinope	2	3	1	3	1	1	1	1	1	2	2	1	2	3	1	1	1	2
pres92	long	random	4	2	4	3	3	2	3	1	2	3	3	3	2	3	4	2	1	5
pres92	long	plal	3	2	5	4	5	4	3	4	5	4	5	3	4	4	5	3	4	5
pres92	long	copern	3	3	5	2	2	3	1	4	5	3	4	2	1	3	4	2	4	4
civil	short	random	2	1	2	1	1	2	1	2	1	1	1	2	2	3	1	1	1	2
civil	short	init	3	3	4	4	4	3	1	3	4	3	5	3	1	1	5	3	2	4
civil	short	copern	4	3	5	3	3	2	1	5	3	5	3	2	2	3	2	3	3	3
civil	short	lsa	2	2	3	2	1	2	1	2	2	3	1	1	1	3	2	2	2	3
civil	short	word	2	1	2	2	1	2	2	3	2	2	1	2	2	3	2	2	1	2
civil	short	sinope	1	1	3	2	1	3	2	3	4	2	1	2	2	2	1	2	1	2
civil	short	nolsa	2	2	4	1	1	2	1	2	2	1	2	4	2	3	1	2	1	3
civil	short	plal	3	3	5	2	2	2	1	2	3	3	4	2	2	3	4	3	3	3
civil	long	lsa	2	3	5	3	2	4	1	4	2	3	3	3	3	4	3	4	4	4
civil	long	word	2	2	3	2	2	3	2	4	4	2	3	4	4	2	4	2	3	3
civil	long	random	2	3	3	3	1	3	2	5	3	2	2	3	3	3	1	3	2	4
civil	long	plal	3	4	5	4	2	4	1	4	4	4	5	3	2	4	5	4	3	5
civil	long	nolsa	2	3	5	2	3	2	1	4	5	3	3	3	2	3	4	3	4	4
civil	long	sinope	3	2	2	3	2	4	2	3	3	1	3	2	2	2	3	3	4	3
civil	long	copern	4	3	4	2	2	3	1	5	4	3	3	3	3	4	3	4	4	4
civil	long	init	3	3	3	4	2	3	2	3	3	3	4	2	3	3	3	3	2	3

Bibliography

B. Baldwin, J. C. Reynar, M. Collins, J. Eisner, A. Ratnaparkhi, J. Rosenzweig, A. Sarkar, and B. Srinivas. University of Pennsylvania: Description of the University of Pennsylvania system used for MUC-6. In *Proceedings of the 6th Message Understanding Conference*, October 1995.

P. B. Baxendale. Machine-made index for technical literature—an experiment. *IBM Journal of Research and Development*, 2:354–361, 1958.

Meru Brunn, Yllias Chali, and Barbara Dufour. UofL summarizer at DUC 2002. In *Proceedings of the Workshop on Automatic Summarization, 40th Annual Meeting of the Association for Computational Linguistics*, volume 2, pages 39–44, July 2002.

Jaime G. Carbonell and Jade Goldstein. The use of MMR, diversity-based reranking for reordering documents and producing summaries. In *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 335–336, August 1998.

Carp Technologies. Sinope Summarizer, version 1.1. Computer software, 2001. Website: `http://www.sinope.nl/en/sinope/index.html`.

W. K. Chan, T. B. Y. Lai, W. J. Gao, and B. K. T'sou. Mining discourse markers for Chinese textual summarization. In *Proceedings of the Workshop on Automatic Summarization, 38th Annual Meeting of the Association for Computational Linguistics*, pages 11–20, 2000.

Freddy Choi. Advances in domain-independent linear text segmentation. In *Proceedings of the 1st Conference of the North American Chapter of the Association for Computational Linguistics and the 6th ACL Conference on Applied Natural Language Processing*, pages 26–33, April 2000.

Columbia Encyclopedia. English civil war. In *The Columbia Encyclopedia*. Columbia University Press, sixth edition, 2001a. Available online at `http://www.bartleby.com/65/en/EnglshCW.html`.

Columbia Encyclopedia. Kazkhstan. In *The Columbia Encyclopedia*. Columbia University Press, sixth edition, 2001b. Available online at `http://www.bartleby.com/65/ka/Kazakhstan.html`.

Copernic Technologies. Copernic Summarizer, version 2.0. Computer software, December 2001. Website: `http://www.copernic.com/desktop/products/summarizer/`.

Copernic Technologies. Copernic summarization technologies. White paper, Copernic Technologies, April 2002. Available online at `http://www.copernic.com/data/pdf/summarization-whitepaper-eng.pdf`.

Hamish Cunningham, Diana Maynard, Kalina Bontcheva, and Velentin Tablan. GATE: an architecture for development of robust HLT applications. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, 2002.

Scott Deerwester, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman. Indexing by latent semantic analysis. *Journal of the American Society For Information Science*, 41:391–407, 1990.

H. P. Edmundson. New methods in automatic abstracting. *Journal of the Association for Computing Machinery*, 16(2):264–285, 1969.

George W. Furnas, Thomas K. Landauer, L. M. Gomez, and Susan T. Dumais. Statistical semantics: Analysis of the potential performance of key-word information systems. *Bell System Technical Journal*, 62(6):1753–1806, 1983.

Gregory Grefenstette and Pasi Tapanainen. What is a word, what is a sentence? Problems of tokenization. In *Proceedings of the 3rd Conference on Computational Lexicography and Text Research*, July 1994.

Vasileios Hatzivassiloglou, Judith L. Klavans, Melissa L. Holcombe, Regina Barzilay, Min-Yen Kan, and Kathleen R. McKeown. SimFinder: a flexible clustering tool for summarization. In *Proceedings of the Workshop on Automatic Summarization, 2nd Conference of the North American Chapter of the Association for Computational Linguistics*, 2001.

Marti Hearst. TextTiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics*, 23(1):33–64, 1997.

J. Hobbs. On the coherence and structure of discourse. Report CSLI-85-37, Center for the Study of Language and Information, Stanford University, 1985.

J. W. Irwin. The effect of linguistic cohesion on prose comprehension. *Journal of Reading Behaviour*, 12(4):325–332, 1980.

Murat Karamuftuoglu. An approach to summarisation based on lexical bonds. In *Proceedings of the Workshop on Automatic Summarization, 40th Annual Meeting of the Association for Computational Linguistics*, volume 2, pages 86–89, July 2002.

April Kontostathis and William M. Pottenger. A mathematical view of latent semantic indexing: Tracing term co-occurrences. Technical Report LU-CSE-02-006, Computer Science and Engineering Department, Lehigh University, 2002.

Partha Lal and Stefan Rüger. Extract-based summarization with simplification. In *Proceedings of the Workshop on Automatic Summarization, 40th Annual Meeting of the Association for Computational Linguistics*, volume 2, pages 90–96, July 2002. Project website: `http://www.doc.ic.ac.uk/~srueger/pr-p.lal-2002/`.

Thomas K. Landauer, Peter W. Foltz, and Darrell Laham. An introduction to latent semantic analysis. *Discourse Processes*, 25(2&3):259–284, 1998.

Steve Lawrence. Online or invisible? *Nature*, 411(6837):521, May 2001.

D. H. Lie. Sumatra: a system for automatic summary generation. In *Proceedings of the 14th Twente Workshop on Language Technology*, December 1998.

Chin-Yu Lin and Eduard Hovy. Manual and automatic evaluation of summaries. In *Proceedings of the Workshop on Automatic Summarization, 40th Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 45–51, July 2002.

R. Longacre. The paragraph as a grammatical unit. In T. Givon, editor, *Syntax and Semantics, 12*. Academic Press, 1979.

H. P. Luhn. The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2(2):159–165, 1958.

Inderjeet Mani. *Automatic Summarization*. John Benjamins, Philadelphia, 2001.

William C. Mann and Sandra A. Thompson. Rhetorical structure theory: A framework for the analysis of texts. Technical Report ISI/RS-87-185, Information Sciences Institute, University of Southern California, 1987.

William C. Mann and Sandra A. Thompson. Rhetorical structure theory: Towards a functional theory of text organization. *Text*, 8(3):243–281, 1988.

Daniel Marcu. From discourse structures to text summaries. In *Proceedings of the Workshop on Intelligent Scalable Text Summarization, 35th Annual Meeting of the*

Association for Computational Linguistics and the 8th Conference of the European Chapter of the Association for Computational Linguistics, pages 82–88, July 1997.

Daniel Marcu. Discourse trees are good indicators of importance in text. In Inderjeet Mani and Mark T. Maybury, editors, *Advances in Automatic Text Summarization*, pages 123–136. MIT Press, Cambridge, 1999.

Daniel Marcu. *The Theory and Practice of Discourse Parsing and Summarization.* MIT Press, November 2000.

Daniel Marcu and Abdessamad Echihabi. An unsupervised approach to recognizing discourse relations. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pages 368–375, July 2002.

Microsoft Corporation. Microsoft Word X for Mac. Computer software, 2002. Website: `http://www.microsoft.com/mac/wordx/`.

Tristan Miller. Essay assessment with latent semantic analysis. *Journal of Educational Computing Research*, 28, 2003.

Douglas C. Montgomery. *Design and Analysis of Experiments.* John Wiley & Sons, fifth edition, 2000.

Jane Morris and Graeme Hirst. Lexical cohesion, the thesaurus, and the structure of text. *Computational Linguistics*, 17(1):21–48, 1991.

Alan Murray. Democratic candidates and the issues. *Wall Street Journal*, page A1, 6 January 1992.

David Nadeau and Nicole Tourigny. Évaluation d'un système pour le résumé automatique de documents électroniques. In *Proceedings of the 14th Canadian Conference on Artificial Intelligence*, pages 277–286, 2001.

David D. Palmer and Marti A. Hearst. Adaptive sentence boundary disambiguation. In *Proceedings of the 4th ACL Conference on Applied Natural Language Processing*, October 1994.

Anthony Pratkanis and Elliot Aronson. *Age of Propaganda: The Everyday Use and Abuse of Persuasion*. W. H. Freeman and Company, New York, revised edition, 2001.

Jeffrey Reynar and Adwait Ratnaparkhi. A maximum entropy approach to identifying sentence boundaries. In *Proceedings of the 5th ACL Conference on Applied Natural Language Processing*, 1997.

Gerard Salton and M. J. McGill. *Introduction to Modern Information Retrieval*. McGraw–Hill, New York, 1983.

Gerard Salton, Amit Singhal, Mandar Mitra, and Chris Buckley. Automatic text structuring and summarization. *Information Processing and Management*, 33(2):193–207, 1997.

Bart Selman and Henry Kautz. Noise strategies for improving local search. In *Proceedings of the 12th National Conference on Artificial Intelligence*, pages 337–343, 1994.

Karen Spärck Jones. A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1):11–20, 1972.

Karen Spärck Jones and Julia Rose Galliers. *Evaluating Natural Language Processing Systems*. Lecture Notes in Artificial Intelligence 1083. Springer, Berlin, 1996.

P. Turney. Learning algorithms for keyphrase extraction. *Information Retrieval*, 2(4): 303–336, 2000.

T. A. van Dijk. *News as Discourse*. Lawrence Erlbaum, 1988.

Vladimir N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, 1995.

Michael White and Claire Cardie. Selecting sentences for multidocument summaries using randomized local search. In *Proceedings of the Workshop on Automatic Summarization, 40th Annual Meeting of the Association for Computational Linguistics*, volume 1, pages 9–18, July 2002.

Michael White, Claire Cardie, Vincent Ng, and Daryl McCullough. Detecting discrepancies in numeric estimates using multidocument hypertext summaries. In *Proceedings of the 2nd International Conference on Human Language Technology Research*, 2002.