Focused and aggregated search: a perspective from natural language generation

Cécile Paris · Stephen Wan · Paul Thomas

Abstract Users are often faced with complex information needs that are not easily represented as a single query. With current technology, the burden of issuing these individual queries, analysing retrieved documents for relevance, as well as aggregating results falls upon the time-poor and informationally overloaded user. Aggregated search techniques represent the new generation of search applications that endeavour to help users perform these complex tasks. However, the way in which different data types are combined in current aggregated search applications is often performed using static hard-coded structures. We suggest that a useful alternative is to marry techniques from natural language generation, such as text planning and summarisation, in order to dynamically determine the best organisation of retrieved information. These organisations can be motivated by linguistic theories that consider issues such as the role that the information plays to facilitate a task, and the relationships between different pieces of information. With reference to a discourse strategy, it is possible to draw on several data sources automatically to generate a useful, focused, and coherent answer. We focus on exploring the parallels between aggregated search and natural language generation in the hope that the fields can be mutually informed, leading to further advances in the way search technologies can better serve the user. These issues are discussed and presented with examples of existing systems across different domains.

Keywords Orchestration · Discourse strategy · Linguistics

C. Paris · S. Wan
CSIRO ICT Centre, PO Box 76, Epping, NSW 1710, Australia
e-mail: cecile.paris@csiro.au

S. Wan
e-mail: stephen.wan@csiro.au

P. Thomas (✉)
CSIRO ICT Centre, GPO Box 664, Canberra, ACT 2601, Australia
e-mail: paul.thomas@csiro.au

1 Introduction

To work effectively in information-rich environments, users must be able to distill the most appropriate information from the deluge of available data. Often, information seeking tasks require a number of related queries, and users must have tools that help them obtain appropriate information from an increasingly diverse range of sources—including the web, content management systems, databases, email, and RSS feeds, to name but a few—and deliver that information in a comprehensible manner.

Search technologies are able to retrieve relevant documents, or relevant passages within documents but, until the introduction of aggregated search, they were unable to integrate the retrieved information into a single result. Furthermore, they are still typically unable to integrate the results of the *multiple* searches required to satisfy a user's complex information need into a coherent whole, in a form that is readable and understandable. Yet, in many situations, people have such complex information needs, requiring access to several information sources and multiple queries. Metasearch engines (Callan 2000) are starting to facilitate this task by providing a single interface to multiple sources, but the user is responsible for issuing several queries to satisfy one complex need. For example, someone planning a trip will require information that is typically found in a travel guide, including information about flights, accommodation and excursions. They may also potentially be interested in related information such as general information about the country or region. With current technology, the burden of issuing the individual queries to obtain each separate piece of information (e.g., general information, flights, hotels, etc), of analysing the retrieved documents for relevance, as well as of aggregating results falls upon the time-poor and informationally overloaded user.

An important research question in natural language generation (NLG) has been exactly the aggregation of information into a coherent whole. We suggest here that NLG might bring a different perspective to the problems being addressed by focused and aggregated search: it examines how the user can be provided with more direct access to relevant information, by retrieving and aggregating all relevant information and combining it into a coherent whole. We refer to the result of this process as an "answer space". This answer space might include specifically all the data items (textual fragments, images, etc) deemed relevant to answer the information need (in which case we might consider the result as a single "answer document"), or it might include links into the underlying information space, in recognition of the fact that there might not be a single answer. We believe that this perspective complements the current approaches in information retrieval and that combining the approaches would be fruitful, especially when dealing with complex information needs. This is what we explore in this discussion article.

The article is structured as follows. In Section 2, we first describe the information seeking tasks which we are particularly interested in supporting. Section 3 examines how researchers from information retrieval (IR) and NLG approach the problem of answering an information need, describing the NLG approach at more length. Section 4 discusses the concept of an "answer space", and explores how it might be constructed, exploiting both IR and NLG techniques. In Section 5, we expand on the NLG approach, providing some work that can be viewed as aggregated search, though from a different perspective. We describe, in Section 6, three general classes of approaches that can be used to implement a system that provides a coherent answer space: a top-down approach, a bottom-up approach and a hybrid approach. We then discuss future work and directions.

2 Background

We are interested in answering complex information seeking needs: that is, information needs not addressed by a single fact, web link or data source, but rather through the integration of information from multiple, potentially heterogeneous, sources. Each source may have its own access interface—for example, databases have one means of access; public websites another; and email yet another—and there is commonly no standardised interface between these. The tasks we are interested in thus require a number of related queries, issued to separate services and eliciting different types of response. Obtaining information to plan for a trip is an example of such complex information needs. Other examples, in an enterprise context, might include finding out about a company to prepare for an introductory meeting: one would need to know what a company does and their physical address, and also potentially information about their standing in the stock exchange, whether they have been in the news in the recent past, whether one's own organisation has already had contracts with them, etc. In both these cases, a user would need to issue a number of different queries and trawl through each set of results to find the appropriate answer(s). In the enterprise context, the queries would need to access different sources of information (both external and internal to an organisation).

Currently, the burden of managing these interfaces and issuing individual queries typically falls on the users. The responsibility for analysing the results also rests with the users. They need to identify the relevant returned documents, find the relevant portions within these documents, and consolidate disparate pieces of information. Having to specify manually various queries and then recognise which portions of a returned document are relevant introduces a tedium that is endemic to scenarios with complex information needs, one that runs the risk of aggravating the often time-poor user.

To address these needs and to support the underlying task of the user, we explore mechanisms that automate parts of the process and streamline the retrieval of relevant information. Some systems already cater to the user in this regard—for example, see Xu and Croft (1996), Kamps et al. (2008), Thomas (2008), and references therein. Their focus is on issues such as multiple queries and appropriate granularities of results in their overall design. However, how retrieved results are organised and presented is also a pertinent design consideration. We believe that delivering the information appropriately is as important as retrieving the most relevant information: this is crucial for ensuring that the results (and their potential relationships) are easily understood and for finding the information quickly. This belief is supported by work in information retrieval in which people have been observed to prefer structured results over lists (Wilkinson et al. 2001). Similar findings have been made in natural language processing (e.g. Isard et al. 2003; Colineau and Paris 2007).

We argue that the appropriate organisation of information is an important element of aggregated search. Such organisation can help the user make sense of how each contributing piece of information from a variety of different sources relates to one another. The aim of this paper is to see if insights about such an organisation from NLG, where such issues have traditionally been studied, can inform research on aggregated search systems; and how we could combine methods from both fields to design and implement more effective systems that attempt to satisfy complex information needs.

3 Satisfying an information need: two perspectives

3.1 Information retrieval

In the traditional search paradigm, a user enters a query to indicate an information need. A system then constructs a list of relevant documents. Although some automated assistance may be provided (see for example the work of Hearst (2006); Jones et al. (2006) or Cutrell et al. (2006) and references therein), the onus is on the user to navigate through that list. This involves potentially exploring a large set of results, remembering browsing paths, bookmarking pages on the way, and potentially issuing more queries. The user then needs to gather the useful information that has been found along the way; perhaps cut-and-pasting from the search results. Essentially, the user is still responsible for determining what information is available amongst the results, as well as for determining the relationships between various pieces of information. Importantly, however, these IR approaches have the advantage of being widely applicable, covering a large range of tasks and data types.

In recent information retrieval research, this paradigm has changed with the introduction of aggregated search (Lalmas and Murdock 2008), and faceted search (Hearst 2006), which adds category information to results. Given an information need, the system constructs an answer that is more than simply a list of relevant documents. It might provide, for example, digest pages (Shushmita et al. 2008), or expand the query into several related queries and provide results grouped appropriately. The results page might present text and images about a topic in different on-screen panels; or provide geographical information with a map (Zaragoza 2008). However, in such work, although a range of different data types are retrieved to support the user's underlying information need, results are still often presented as a list (or sets of lists).

We also see a move away from the presentation of results simply as a disjoint list of documents, with the advent of focused search (Trotman et al. 2008), which includes passage and element retrieval and question-answering. Here, systems try to go beyond returning a list of documents. They attempt to provide information that will directly answer the user's need by retrieving the appropriate fragments of text. A system might, for example, extract relevant passages and present them in the context of a table of contents (see Szlavik et al. 2007).

In natural language processing research, multi-document summarisation retrieves and collates similar information at the granularity of sentences. For example, the summarisation of news items on the same topic from various news agencies has been previously explored by McKeown and Radev (1995) and Barzilay and McKeown (2005). Indeed, this line of research has been integrated with work on aggregated search (Shushmita et al. 2008) and focused search, where the retrieved passages are then summarised (Trotman et al. 2008). These approaches all shift the burden of identifying the relationships between returned results away from the user.

3.2 Natural language generation

Work in natural language generation (NLG) has also looked at the problem of satisfying an information need, from the perspective of providing an automatically generated text that answers some underlying question or questions that a user may have, without the user having to do more searching or browsing through documents. In addition, NLG research has focused on constructing a natural response to a question, presenting the right information in an appropriate linguistic form. The natural and coherent response is aimed at

helping the user understand and act upon the answer without a lot of navigation through a large set of results. The NLG work we describe in this paper additionally creates output that is sensitive to the user's profile. For example, a system might consider constraints regarding user preferences for language choice as well as previous user interactions, to name but a few.

This naturalness of information presentation has been explored drawing on linguistic studies of human communication. Researchers have looked at issues such as conversational implicature (what will be inferred naturally by a reader from an answer) to ensure that the answer would be understood appropriately (Grice 1975).

The observation that information needs can be complex and can require more than a single fact (or, with current technology, more than one web link or web resource) to be satisfied is akin to the observation that McKeown (1985) made in her work on producing paragraph-length informative texts in response to an information need. In that work, McKeown argued that some queries are such that people expect their answer to include several facts, organised in some prototypical fashion, referred to as a schema. This organisation is reflected naturally in the way one writes text. For example, to describe an object, one typically defines the object in terms of its superordinate(s) (e.g., *a stool is a piece of furniture*), provides its components and/or its attributes, compares it to a known object or differentiates it from another similar object (e.g., *a stool is like a chair except that*...), and finally gives a concrete example of it (McKeown 1985). As a result, in requesting a description of an object, one expects an answer comprising these facts, organised in this manner. Such an organisation highlights the relationship between the different types of information to be presented in terms of the function they play in the information exchange.

Other types of prototypical structures exist. For instance, when presenting a sequence of events, we normally present them in chronological order. Using terms from a linguistic theory of coherence, we could say that each fact presented is related to the previous one by a *sequence* relationship. Or to explain how a process works, one usually presents *cause-effect* relationships (i.e., this happens causing this other thing to happen) (Paris 1987). As a further example, when explaining something to someone, we usually provide *background* or *contextual* information first.

The notion of a prototypical organisation for information applies more widely. For example, the way one engages in a conversational dialogue and structures the delivery of ideas also follows certain patterns (Moore and Paris 1993). Similarly, in terms of current technologies, the production of multimedia documents also arranges data of different media in a certain way (e.g., see Bateman et al. 2001; Wahlster et al. 1993). Because of the way people normally communicate, one tends to juxtapose one piece of information with another for deliberate reasons. That is, the organisation of a text is not random. Producing text that conforms to these expected patterns of communication thus helps the reader make sense of the information (and avoids miscommunication).

Schemas, as representations of such prototypical structures are static, however, and have limited flexibility. Other approaches allow this organisation to be dynamic depending on the changing information need and the availability of related content—e.g., see Moore and Paris (1993). These approaches exploit planning mechanisms (Sacerdoti 1977), and linguistic theories of coherence—e.g., rhetorical structure theory (RST) (Mann and Thompson 1988). They work as follows: the system is given a communicative goal that is meant to satisfy the user's information need, itself determined by a query input processor. As the result of planning this goal, an answer (for example, the generated document) is constructed.

The top-level goal gets decomposed into sub-goals, which are in turn decomposed, until the recursion bottoms out at leaf-level pieces of information called *primitive speech acts* (Austin 1962). At that point, there is both a very specific communicative goal and a context enabling a specific sub-query to be executed to retrieve information from the underlying knowledge sources in order to provide the content needed to instantiate these speech acts. The knowledge can be represented in a number of ways, including databases, symbolic knowledge bases, or XML data. Personal information about the user is also often included to tailor the query. The results of the particular sub-query can be easily inserted into an automatically determined structure, where it plays a particular communicative goal. We describe this approach in more detail through examples in the next section.

From an NLG perspective, these prototypical ways of organising information are referred to as "discourse strategies". Essentially, these embed the answers to information needs within an appropriate context, helping the user to interpret and act on that knowledge. The strength of NLG approaches is that they alleviate the problem of making sense of a disparate list of results, since the answer document delivers appropriate and relevant information in a coherent overall structure. They also avoid the burden of issuing many queries, by decomposing the high-level communicative goal that is believed to answer the user's information need and automatically retrieving information as appropriate.

These strategies are explicitly represented within a system and reasoned about by that system, dictating which information is to be retrieved, and how it is to be aggregated. Typically, these strategies need to be specified in advance as a planning resource and are the result of linguistic analyses. While some strategies have already been identified by linguists and computational linguists and some are intended to be domain-independent, a new application requires additional analysis to define extra strategies for specific domains. Recent work has explored ways in which the manual cost of this specification can be reduced via corpus-based methods, exploiting theories of coherence and cohesion and machine learning (e.g., Barzilay and Elhadad 1997; Marcu 2000; Barzilay and Lee 2004; Wan et al. 2008; Sauper and Barzilay 2009).

NLG applications usually address a limited range of information needs, compared to search technologies which can respond to arbitrary under-specified queries. NLG systems typically consider a particular application domain and focus on those application-specific needs which have been identified beforehand. A focused application is not necessarily a weakness. However, such approaches typically incur the additional cost of understanding how these needs can be satisfied, and the set of underlying information sources must be known. We suggest in this paper that, as researchers are developing automated ways of acquiring rhetorical information, discovering relationships between data items (or clusters of items) and understanding what is expected to satisfy an information need, NLG techniques coupled with IR techniques might be useful in a larger range of situations.

3.3 A continuum of options

We can characterise IR and NLG as defining the two ends of a continuum based on the genericity of the approaches, the appropriateness to task and context, and the understandability and naturalness of the results, as shown in Fig. 1. We see aggregated search as being closer to the NLG end of the continuum. Both must identify exactly which information addresses the user's needs and both must additionally understand how to aggregate and present the results. There must also be some understanding of the range of retrievable information (e.g., one must know that all retrieved paragraphs are news items).
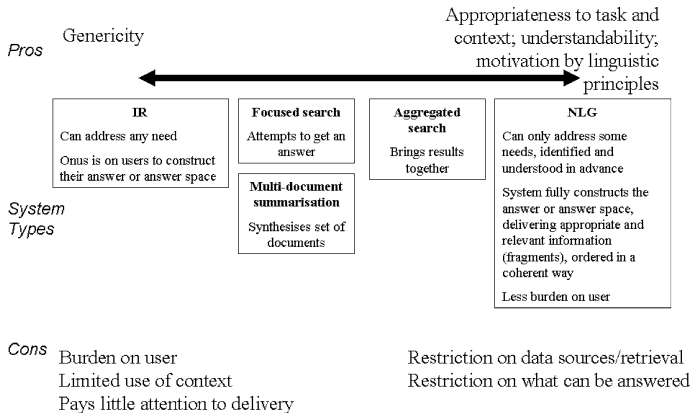
Fig. 1 Different ways to satisfy an information need

We see focused search as tending to lie closer to the IR end of the continuum. This includes situations where multi-document summarisation is employed to provide a synthesis of the retrieved passages. Approaches are typically generic (through their use of statistical techniques), and do not require knowledge about the underlying knowledge sources. At the furthest extreme, traditional search engines can respond to any query but do little to organise the results past listing entire documents in order.

Importantly, genericity of systems need not be at the sacrifice of information structure. Focused search approaches that include multi-document summarisation and faceted search are testament to this: both employ generic approaches, and yet both are motivated by the linguistic principles of relevance, cohesion and coherence through the use of additional structure.

We postulate that both research in NLG and aggregated search can usefully complement and benefit from each other. There is already some conceptual overlap. The notion of a relationship between information items is related to the concept of *affordance* which is used in the fields of IR and human computer interaction (HCI): that is, a data type can lend itself to certain presentation styles. For example, time is well represented through something linear, locations through the visual representation of maps, and a text passage from a large structured document can be presented in the context of the table of contents of the document from which it is retrieved. We can summarise some lessons learnt from NLG that might be applicable to aggregated search:

– There are some prototypical ways of organising information that are natural to human readers. These are often called "discourse strategies"; when these are explicitly represented, additional reasoning about how to present information can be performed; we argue that one reason why aggregated search is useful is because it embodies (albeit implicitly) some such patterns.

– Making the relationship between data items explicit can help people understand information. In text, for example, one uses "cue phrases" to indicate the relationship between sentences of clauses: "for example" explicitly tells the reader that what follows is an example of what was just introduced; "as a result" makes explicit the causal relationship between two events, statements or facts; "however" indicates a contrast. When one considers layout as well, some typographical devices help make

these relationships explicit: an itemised list can indicate a series of parallel items; an image or a table will have its caption (the statement introducing the image or table) just below or above it; etc. Some relationships are obvious and will simply be assumed (e.g., the text immediately underneath a figure will be taken to be its caption). Providing a system with these explicit representations also enables additional reasoning.

– Linguistic principles of coherence and cohesion can help aggregate a set of facts in such as way as to be useful to the user. Aggregation can be at a high level (discourse level, like schemas and discourse plans), or a more local notion of coherence (e.g., at the clause level, bringing two facts together because one expands on the other, as is often done in multi-document summarisation).

What aggregated search and NLG do, in providing a response to an information need, is to present the range of useful data types that, together, can answer this need (i.e., each individual item contributes to the need in some way). We call this range of answers an "*answer space*". We now introduce this concept more fully and explore how we can exploit techniques from IR and NLG to construct such a result.

4 Defining an answer space

The idea of an "answer space" recognises that, in some cases, it might not be possible to extract exactly the relevant data items required to answer an information need: there might not be a single answer, but rather a set of aggregated information items with a large information space behind it that the user might want to navigate (Teevan et al. 2004). We distinguish an answer space from a list of results through its structure. An answer space has a coherent structure (whether implicit or explicit); it is a restricted set of data, possibly from various sources, arranged in a way that allows the reader to navigate through the space naturally and effectively. Take for example, a brochure delivered over the web, which includes web links. The user can follow any link. However the brochure, as an answer space, indicates how information, including that which is found at the end of a hyperlink, is related. Outputs of aggregated search usually represent an answer space, although the presentation may be limited to a finite number of data items.

Borrowing the notion of natural pattern of communication from NLG, we note that an answer space might include more than just the information deemed "on topic". It might include related and supporting information (e.g., in NLG terms, background or elaborations), and justifications for the content provided. The former might fulfil the users' information need in a more complete fashion, thus pre-empting the requirement to issue more queries. The latter might help readers understand why some information is worth examining. It might also increase the users' trust in the information (or allow them to decide whether the information is trustworthy, by considering the reliability of its source, for example).

The underlying data elements in the answer space can be complex and varied. They can be heterogeneous, ranging from extracted text passages from multiple documents of differing domains and genres, to database entries or meta-data tagged graphics, to mention but a subset. Importantly, these data items are organised in a meaningful and useful way, thus helping the reader make sense of the complex answer space. We suggest that discourse strategies, as used in NLG, are mechanisms to provide such organisational principles.

5 The benefits of a natural language generation approach

Before further examining how work in information retrieval and natural language generation can complement each other, we describe in more detail how NLG methods can aggregate search results and construct an answer space.

We provide three examples of how a system based on natural language generation techniques would produce an answer to satisfy a user's information need. Our first example is in the surveillance domain, the second in the tourism domain, while our third example is a system that produces brochures about an organisation. In these examples, the discourse strategies are represented as plans (and not schemas), and a "text planner" (Moore and Paris 1993) was employed. The examples are taken from applications we have developed (Paris et al. 2001, 2005; Colineau et al. 2004a; Paris and Colineau 2006).

The three examples also serve to illustrate three attributes of an NLG approach. Firstly, the way in which information is structured can be determined dynamically. Secondly, a dynamic structure can be tailored to run-time user query constraints. Finally, user evaluations have revealed results that are favourable to dynamically determined structure, illustrating the benefits of organising the retrieved information. In our descriptions below, we focus on the discourse planning aspect of the systems, as opposed to sentence planning and presentation. This is for two reasons: first, we believe it is this aspect that is most relevant to the idea we are exploring in this paper—i.e., that notions of coherence can help aggregated search systems ensure results are presented in the most useful ways. Second, the specific applications presented below have very limited sentence planning mechanisms, as they access textual data (e.g., text fragments) as opposed to symbolic knowledge bases. Our applications do perform some sophisticated presentation planning (especially when they deal with graphical user interfaces or to reason about space constraints), but this is beyond the scope of this paper.

5.1 Determining a coherent structure: the surveillance report example

We first consider a specific application in a shipping surveillance domain (Paris et al. 2005). Surveillance operators often monitor situations by polling information from a number of different sources, accessing each of them with a different interface. To help them with their task, one could simply collate information acquired from multiple sources. However, a coherent view of a shipping situation could help operators prioritise different elements of the interface as the context changes. Using a discourse planning approach, we developed a system capable of performing aggregation[1] of information tailored to various contexts and tasks. Our application obtains data from various public-domain information sources (e.g., the International Chamber of Commerce Commercial Crime Services, anti-shipping activity messages, detention ship lists, ship movement databases for a port, and weather reports for specific regions) and delivers its output as a multi-page website for port surveillance operators. A sample output is shown in Fig. 2.

Based on an analysis of the domain and of users' tasks (as is typical in human computer interaction—see Diaper and Stanton 2004), we identified that the important part of such a report is the information about high-risk ships in the region and piracy attacks that have

---

[1] Note that "aggregation" here is different from the typical use of the term in Natural Language Processing. There, the term "aggregation" usually refers to a semantic or lexical operation. For example, instead of listing all the days of the week (Monday, Tuesday, etc.), one could perform an "aggregation" and say: "every week day" (Shaw 1998).
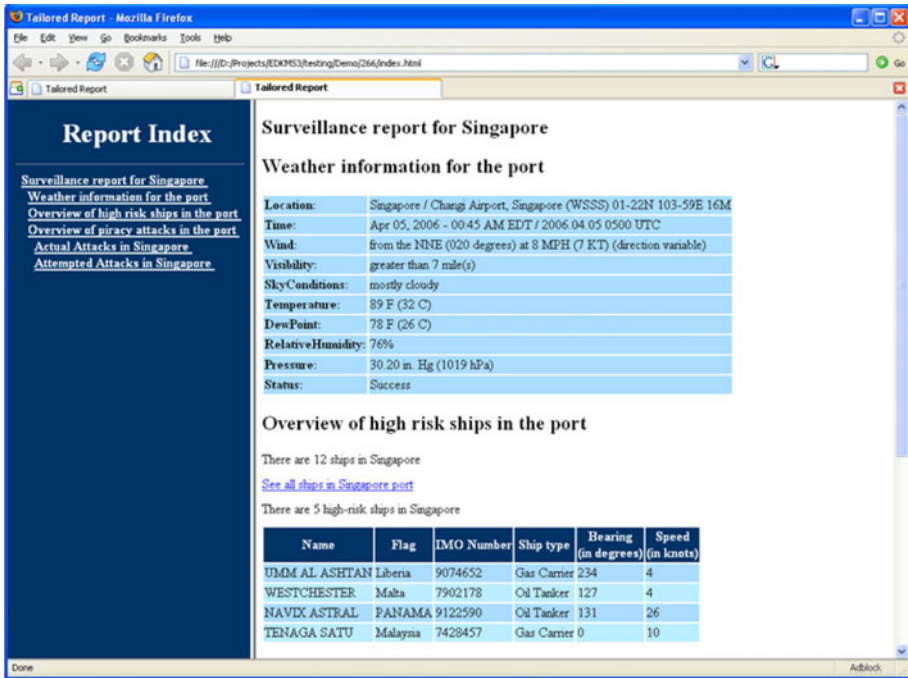
Fig. 2 A report for a surveillance operator

Fig. 3 A discourse strategy to produce a report for a surveillance operator

Discourse Strategy for a Report for a Surveillance Operator
    Communicative goal to be achieved:
        Provides a report for a port city for a surveillance operator
    Discourse Strategy for this goal
        Preparation: Provide a title
        Context: Give the weather report for port city
        Provide report for ship for port city
        Provide report for piracy attacks for port city

occurred (Colineau and Paris 2003; Colineau et al. 2004a). While this might be the most important information, the weather report for the region is also important (because it might have a potential impact on the high-risk ships), and it should be provided as the context in which to interpret the other information items. The text planner starts with a discourse strategy that plans the overall structure for a surveillance report, as shown in Fig. 3. This discourse strategy is the one employed to produce a report such as the one of Fig. 2. It indicates how to decompose the communicative goal and the coherence relationships between the information to be retrieved (through sub-goals). In the figure, for example, the discourse strategy indicates that a title would be in "*preparation*" to the core information of the report; similarly, giving the weather details provides "*context*". In the system, these strategies are embodied in separate plans. (The number of plans required depends on the granularity of the information and the reasoning that must be performed. In this application, as we needed fine-grained reasoning about how to present the information, we had 49 discourse plan operators (Paris et al. 2009).) At run-time, the top-level goal of producing a surveillance report for a specific region is posted. It is decomposed through the planning process into sub-goals according to the plans.
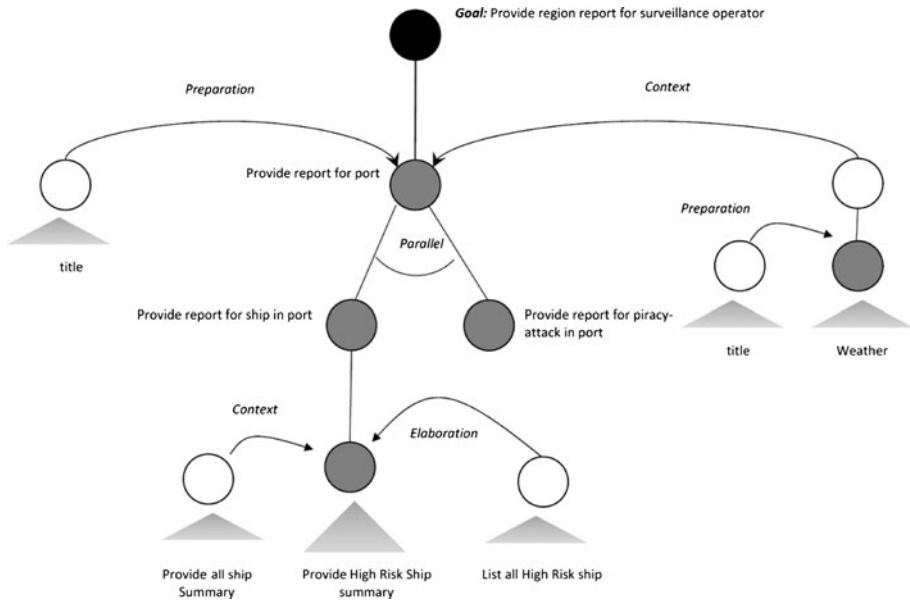
Fig. 4 (Partial) discourse structure of the information displayed in Fig. 2

The planning process produces a tree structure, often called a "*discourse tree*" or "*discourse structure*" (Moore and Paris 1993; Colineau et al. 2004b; Paris et al. 2009). This is illustrated in Fig. 4: sibling nodes have explicit annotations encoding the rhetorical relationship between the siblings' information. Consider again the report in Fig. 2, with its discourse strategy of Fig. 3. The top level goal was decomposed into two main goals (called "nuclei", in RST (Mann and Thompson 1988), the specific linguistic theory we are exploiting), indicated with the grey nodes in the middle of the tree shown in Fig. 4. Both of these constitute the main aspect of the report, and they are in a *parallel* relationship to each other. In contrast, the sub-goals that will lead to producing the title and the weather details are supportive information (called "satellites" in RST). The tree explicitly indicates their relationship to the main parts of the report, as marked on the arcs of the tree structure, namely "*preparation*" and "*context*". Each sub-goal can be decomposed further. One such decomposition is shown in the figure. It also shows one additional possible relationship: "*elaboration*", indicating that the node provides additional details. So, when constructing an answer, a discourse tree is built.

Specific requests for information can be formulated to retrieve data. They are placed in context within the overall document structure. This placement within the structure helps provide additional contextual parameters that can inform the data retrieval query. For example, a specific request to a weather service can be issued with the appropriate location to fulfil the goal of providing the weather details for the region. Similarly, a specific request to find all the XML fragments representing piracy attacks in a specific region is sent to a service that accesses an XML file of all the piracy attacks. Requests to access data can be serviced via a variety of information sources. In this application, the system issues requests to data bases and XML files, retrieving the relevant data.

These data sources are heterogeneous: they are all accessed through different interfaces, return different data types and deal with different topics (e.g., weather reports, crime services, ship data). The planning process discussed here allows the system to

automatically generate appropriate queries for each source. When relevant passages are retrieved from the sources and aggregated into a coherent virtual document, they appear semantically coherent.

The discourse structure shows the various communicative goals that were posted, and their relationships in terms of coherence relations. We see in this example that, using this approach, a system can produce answer documents which are not just text—here, tables have been chosen as the best presentation mode, with a summary sentence to introduce them (Presentation decisions are done through *presentation operators* (Paris et al. 2009)). The results of the specific queries issued were combined, or aggregated, based on the underlying planning strategy. These results can be sourced from a range of different data types. In particular, textual data can be retrieved at the granularity of passages and sentences, thus illustrating the connection to focused search.

There can be a number of competing planning strategies and therefore a number of possible plans that are chosen for further decomposition. The system can choose which plan is most appropriate at run-time based on a variety of constraints and thus flexibly respond to changing user needs. In this example, constraints include the region under consideration and its geographical type (whether it is a city, country, etc.), and the user's role (e.g., a surveillance operator or a strategist). They can also consider existing knowledge that the user might have gained from previously generated reports, referred to as the "*discourse history*", since this can be used to highlight changes in the situation (or it can highlight what has not changed, as this might be important in this domain—to produce text such as "this high risk ship is still in the port") (Paris et al. 2009). When several operators that can achieve the same goal and all have their constraints satisfied, the system picks the first operator. (In general, it is possible to add additional heuristics, for example, to try to produce the shortest/longest amount of text, or text most likely to be understood by the user, but this was not done in this application. The interested reader is referred to Moore and Paris (1993) for a description of how this can be done. Moore and Paris (1993), Colineau et al. (2004b) and Paris et al. (2009) also provide detailed descriptions on text planning and how competing operators can be selected, our specific document planner and the applications mentioned here.) A final interesting aspect about this system was that there was no explicit query required on the part of the user: it was triggered by a map-based interface: upon a click, a surveillance report about the region in the focused zone would be generated. A user profile was accessed when the user first logged into the system.

5.2 Incorporating run-time search constraints: the travel information example

In the surveillance example, the user's information need was implicit and governed by the user's interaction with a map-based application. We now present an example of a system where the user can provide additional explicit constraints to tailor the generated output. This is in the context of obtaining information to plan a trip.

A search interface allows a user to indicate which city is of interest and to provide constraints regarding aspects such as food preferences and budget limits. These run-time user-provided constraints are similar to those typically found in focused search interfaces for travel information. The difference is that the response generated by the system again arranges the relevant information in a coherent structure to facilitate ease of reading and to allow the user to easily gauge what information is available.

Using a natural language approach, we can specify the discourse strategies typically employed in a travel guide. We manually performed a discourse analysis of a set of travel guides (based mostly on their table of contents). As for the surveillance application, we
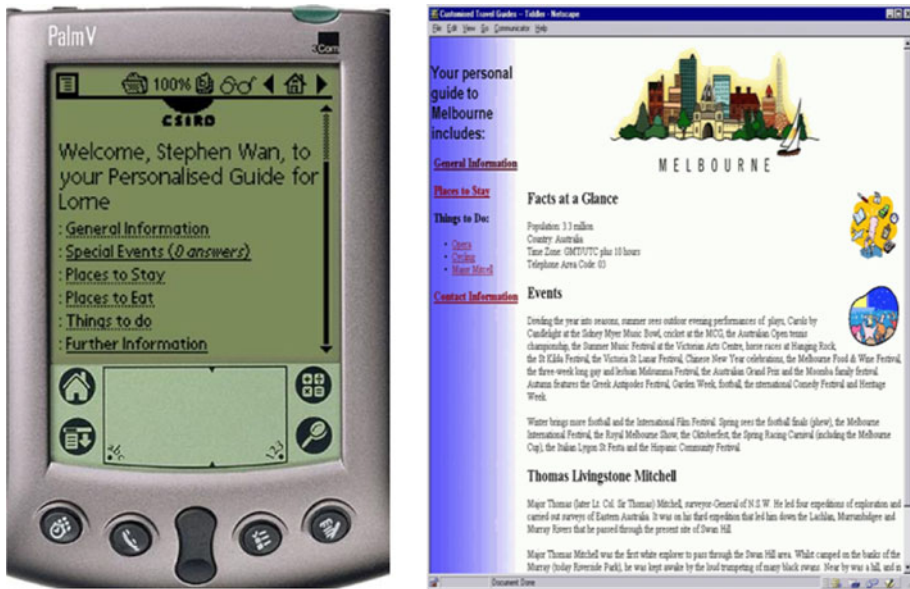
Fig. 5 Query about a travel destination—answer produced by an NLG system

encoded these in plans, which are used by a planning engine to decide what information to include at each point, where to find it, and how to aggregate it. Again, once a piece of information, or a request for that information, has been placed within the overall document structure, the query is sent to the underlying knowledge source and the data retrieved. In this application, returned data is often a text fragment but could be a map or a table of special events. The dynamic structure, sensitive to the run-time constraints entered by the user, aggregates all these separate pieces of data into a coherent whole. This is designed to match the user's expectations of a travel guide.

Figure 5 shows an example of the output to a query about Melbourne. The system has produced a personalised travel guide (Colineau et al. 2001; Paris et al. 2001). All the sections shown (e.g., "General", "Hotels") correspond to elements in a discourse strategy that was defined for this application, again embodied in plans. Through the plans, the system is able to generate a specific query to retrieve the information from the appropriate source (a database of events, or an XML resource). Each automatically created query directly addresses a different part of the information need. The user does not need to issue these separately. Instead, the system plans to include the appropriate data and then retrieves the information.

The way the information is aggregated and presented is dictated by the discourse strategies. Because the system has an internal representation of the discourse strategies describing the coherence relations that link various items of information, it is also able to reason about how much information to present (and what to hide) given specific space constraints (e.g., "real estate" on a mobile phone being smaller than on the web) (Colineau et al. 2001; Paris et al. 2008a, b). For example, Fig. 5 shows two screen shots corresponding to two different displays. Underlying both generated travel brochures is the *same* structure, showing the utility of the discourse tree when a system needs to produce for different delivery devices. The discourse tree enables the system to make presentation-

level decisions to cater for limitations in the devices chosen. However, having the same coherent structure means that the user can swap between devices easily and still maintain the same sense of what information is available and where to find it in the generated document. We performed an early evaluation of this system considering the user's experience when planning an imagined holiday in South-East Australia (Paris et al. 2001). Our evaluation was aimed at establishing the usefulness of the tailoring, and demonstrated some user preference for tailored delivery.

5.3 User evaluations on structure: the 'SciFly' company brochure example

Finally, we briefly describe SciFly, which produces brochures on demand in response to a query from the user (Paris and Colineau 2006). These brochures describe information about an organisation, in this case Australia's science organisation CSIRO. The brochures are tailored to a person's interest(s) and formatted appropriately for paper printing, email delivery or web presentation. The system again makes one pass through the discourse strategy planning phase of the system, as in the case for the tailored travel guide, and the resulting structure provides enough information to later reason about how much to present given the device space constraints (Paris et al. 2008b).

Figure 6 shows a sample two-page PDF brochure produced by SciFly. Because these brochures were meant to replace the manually written ones, they have to be coherent and well presented. In these brochures, the information unfolds in a natural progression of topics, following a structure typical for the genre. This is again declared as a planning strategy. We designed the strategies employed by the system through a manual corpus analysis of the existing brochures and discussions with communication managers to elicit their requirements. (The application employed 50 operators, again because we needed fine-grained control over the ways the brochure was generated, in particular to reason about space constraints.) SciFly accesses an XML database of text fragments and a staff database for contact details. (The XML database is intended to emulate retrieval of information from a content management system. However at the time of implementation, the organisation had not yet migrated fully to the new system, and so an XML database was implemented.) As SciFly retrieves text fragments, there is no need for sentence planning. Staff details are presented in a tabular way ("Phone: ...; Email: ...") rather than as text.

When we first started developing SciFly, we performed a user study to assess whether a tailored hypermedia system (such as SciFly, or its predecessor Percy) offering tailored and coherent information would provide a good alternative to the current search-and-browse facilities, also allowing corporations to be more responsive to their clients' needs (Paris et al. 2003). Specifically, we chose to investigate whether users would prefer a tailored and coherent delivery mechanism over the conventional search engine results. We compared our system with the CSIRO web search engine, Panoptic,[2] which has no NLG facility.

We had twenty participants, selected four topics from the research topics in our institute, adopted a Latin-square experimental design and employed a post-system and an exit questionnaire. The results can be summarised as follows: for questions about content (e.g., "The system provides sufficient information" and "The information provided by the system meets my need") and format ("The structure of the presented information is clear to me", "I think the presented information is organised in a useful format", and "I think the presented information serves well as a useful online brochure"), our system got higher scores than the search system (but the differences were not statistically significant). As for

---

[2] Now called Funnelback: http://funnelback.com/.

Fig. 6 A SciFly brochure

preferences ("The information delivered by the first/second system attracts my attention better", "The first/second system provides a better explanation on why a piece of information is presented" and "Overall, I prefer to use the first/second system as an online brochure of the searched topic"), our system was rated significantly higher. We thus showed that driving the retrieval process using text planning leads to organised generated documents that were favoured by participants over a regular search (Paris et al. 2003).

The study also suggested that producing coherent and tailored hypermedia was an effective way to deliver information about an organisation, probably more effective than the traditional search-and-browse mechanisms. This illustrates the benefits to the user of employing NLG approaches to structuring and aggregating data.

5.4 Discussion

We argue that the systems presented in this section (the surveillance application, the personalised travel guide and SciFly) can be seen as performing focused and aggregated search: focused, because they aim to retrieve information that will directly address the user's needs (not just point to it), thus constructing the answer document; and aggregated, as the systems clearly bring together results. However, the design of these systems has concentrated on how information fits together, and their capacity for search was treated as secondary.

We presented the approach often adopted in natural language generation (NLG) to satisfy a user's information needs, illustrated through three examples. The key points are that an NLG system seeks to provide answers to a user's information needs (not just pointers to information sources that are likely to contain the answer or part of the answer). To do so, it issues specific queries to underlying knowledge sources, obtaining the required information. The system then processes the returned data, if required, to ensure its appropriate delivery, aggregating the resulting set of information chunks to conform to some naturally occurring pattern. Altogether, it constructs the answer. As mentioned earlier, the output could be an answer space, with links into the underlying information space; this enables the user to do additional navigating if required.

We have used this approach in a number of applications, as illustrated in the systems presented. In these applications, without our system, the user would have had to issue multiple queries themselves, potentially using different interfaces, and sift through the results to get exactly what they needed. Our NLG-based system automates this process and aggregates the information into one relevant generated response.

In the natural language generation community, many systems were built using such discourse planning approaches to satisfy a user's information needs in a variety of contexts: to enable a system to participate in educational dialogues, or to produce text or multimedia output—e.g., Moore and Swartout (1991), Moore and Paris (1993), Wahlster et al. (1993), Green et al. (1998), De Carolis et al. (1999). In all cases, an important aspect of the output is its coherence, the way the information is presented, sometimes realised through several complementary media (e.g., text and graphics; picture and speech; etc.). Also important is the fact that such systems have an internal representation of the produced output, with knowledge about what each piece of information is meant to communicate and how the various pieces fit together. This representation can be exploited for a variety of reasoning tasks: to enable systems to participate in a dialogue (e.g. Moore and Swartout 1991), to generate appropriate cue phrases to link two spans of text (e.g. Scott and de Souza 1990), or to reason about layout (e.g., Bateman et al. 2001). (For other applications, see Taboada and Mann 2006). Finally, the various organisational patterns employed have a linguistic motivation.

We now examine some of the limitations of these approaches compared to a conventional search engine in information retrieval. First, although this is not essential, some systems built using these approaches rely on symbolic knowledge bases (i.e., a resource with a fine grained representation of knowledge). These knowledge bases enable precise reasoning of meaning to construct an output. As the structures of these knowledge bases

are known, the appropriate tools (including sentence planning mechanisms) can be used to ensure the end result is grammatically correct. However, these knowledge bases are difficult to obtain. This limitation is typically referred to as the knowledge acquisition bottleneck. This is in sharp contrast with the resources available to a search engine.

It is not always the case, however, that one must have a symbolic knowledge base. This was not the case, for example, in the applications presented above. One used public domain databases and RSS feeds, the other a set of XML files, and the last one a database of XML fragments and a conventional database. What is still the case, though, is that the set of available underlying sources is known as well as the structure of these sources, so that the systems "know" where to obtain what information and what to expect as a result of the various sub-queries. This is what enables them to construct documents that look as if they had been manually written. In addition, because of this knowledge, these systems typically do not have to deal with inconsistent or contradictory information.[3]

Finally, while such approaches produce coherent aggregations of information, they typically require an understanding of the required discourse patterns or of the user's tasks. In the applications we presented in this section, for example, we required an analysis of the user's tasks (in the case of the surveillance applications), or a study of what would be appropriate discourse strategies in the domain (e.g., both the tailored guides and the personalised brochure applications), and an understanding (on our part) as to which information was to be found where, so that we could specify the plans with the appropriate specific queries. The result, though, is somewhat akin to a combination of focused and aggregated search. We thus believe that it might be possible to drive the information retrieval process from constructs such as discourse strategies, when dealing with complex information needs for which expected patterns of answers can be discovered.

We believe that the core linguistic ideas of coherent organisation are useful (even if the way they are currently defined in NLG systems does entail some limitations), and that we can capitalise on them with the automated methods studied in related fields of natural language processing. We briefly mention some work that attempts to make use of automatically acquired coherence relationships to satisfy a user's need.

In summarisation work, for example, Wan et al. (2009) looks at satisfying the information needs of biomedical researchers during their task of browsing scientific literature. Based on an analysis of scientific text and citations, we see that citations in a text are in a coherence relationship with respect to the citing document (Teufel and Moens 2002). In their work, Wan et al. (2009) have assumed that a citation usually elaborates on a citing sentence (no matter what its status is with respect to the work of the citing document as a whole). They exploit this implicit *elaboration* relationship between a citing sentence and a cited document to find, in the latter, the relevant and important sentences for the citation at hand, in its context. Because the system extracts individual sentences, its overall result might not be as smooth as a manually authored summary, yet it provides useful information to the researchers. Importantly, the notion of cross-document coherence (as represented by citation links, for example) is widely-applicable for generating summaries in different scenarios (see also, for example, Teufel and Moens 2002; Qazvinian and Radev 2008; Mohammad et al. 2009).

In general, summarisation research has sought to exploit relationships that exist among sentences, while still remaining domain-independent. For example, some methods look for

---

[3] Some systems that deal with argumentation might deal with contradictory information; they are likely, in that case, to have a discourse strategy to present *contrastive* information, as they want to present both sides of the argument.

clusters of similar sentences using linguistically-motivated features, as in the work on lexical chains by Barzilay and Elhadad (1997). These approaches say something about the structural relationships within the text based on content. These relationships might not be explicit (as in the discourse plans or schema introduced earlier), and they might not be as fine-grained. As a result, they may be less precise due to errors introduced in automated text analysis methods, and coherence might not always be guaranteed. However, they are domain-independent. In addition, the goals of summarisation research are related and relevant to focused and aggregated search, as illustrated by the work of Shushmita et al. (2008), as mentioned earlier. The cross-pollination between NLP and IR is not new. For example, summarisation work borrows heavily from IR, using vector-space approaches (e.g., Radev et al. 2003).

We argue that aggregated search and NLG can further inform each other to answer a user's information needs more effectively, leveraging expected patterns of communication (identified either through a manual analysis or an automatic one when appropriate data is available). In particular, deciding what information to juxtapose can be chosen based on linguistic patterns in addition to conventional IR methods. In the next section, we characterise three avenues of research that explore the possible coupling of the two fields.

6 Classes of approaches to generating answer spaces

There are a number of ways an answer space can be constructed, depending on the level of automation required (or feasible) and the importance of the coherence of the result. Figure 7 illustrates three approaches, as they might generate answers for the company overview task of Sect. 2. The answer space can be dynamically planned, top-down, by explicitly decomposing information needs into their components and imposing an organisation (Fig. 7a). The example shows an ordering of four pieces of information (company details, news about a company, financial information, and past contacts) that might be obtained from a schema. Alternatively, it can be obtained in a bottom-up fashion by discovering and analysing relations already present in the data, organising them in typical patterns (Fig. 7b). The example in the figure shows a hypothetical situation with three automatically determined clusters of retrieved content. The relationship between each pair of documents in the cluster is simply one of similarity in content. Each cluster has been labelled with the title of its first document. The three clusters about "widgets", "share information" and a "new deal" are not presented in any particular order. We believe a hybrid approach (Fig. 7c), using both top-down and bottom-up methods, is a likely candidate approach for future such systems. The example shows a schematic organisation of the information similar to that of Fig. 7a. However, each "section" has been populated with content retrieved and clustered in a bottom-up manner.

Finally, the *representation* for the organisational principles (the discourse patterns) can be automatically derived or statically hard-coded; and coherence can occur at a global level (i.e., at the discourse level, as in the schemata or discourse strategies introduced earlier), or at a more local level (e.g., at the sentence level).

We examine these options in turn in Sects. 6.1 to 6.3.

6.1 Top-down approaches: imposing a coherent organisation

The NLG systems in Sect. 5 all used a top-down discourse planning approach, and the organisation was embodied in plans. The organisation of the space was essentially imposed
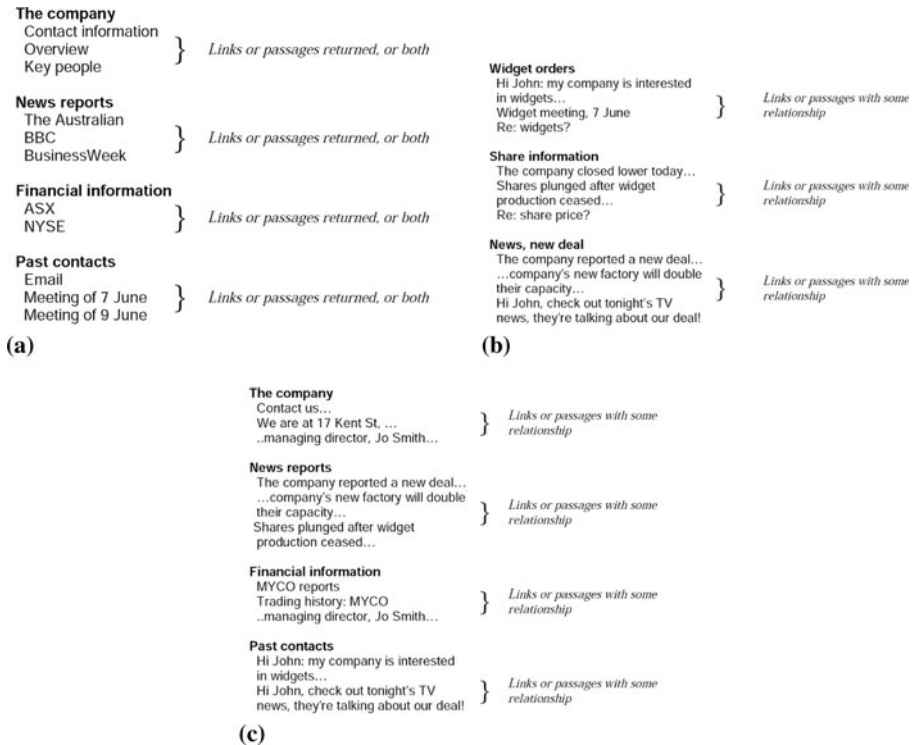
Fig. 7 Three approaches to generating answer spaces, given a complex information need: **a** Top-down: Topics are identified manually or automatically; ordering is coherent; and queries are issued according to discourse strategy. **b** Bottom-up: Queries are issued to all sources, and results analysed to find relationships. **c** Hybrid: A top-down approach directs a set of specific queries; a bottom-up approach then looks for relationships amongst the returned items

on the underlying information sources, and information was returned from very specific queries. In this section, we generalise over the examples presented in Sect. 5 to illustrate how an NLG approach might be adopted in subsequent research.

For top-down approaches, an organisation for the answer space is pre-defined to meet the user's information need. The answer space is then populated with data returned via search engines. We term this the "orchestrator" approach (Paris 2006): the coherent organisation coordinates the issuing of specific queries, and thus the retrieved data will hopefully fit together harmoniously. These queries are conventional.

The imposed organisational structure can be manually coded, based on discourse structures already identified in linguistic or computational linguistic research (for example, biographies have a known structure) or a new corpus analysis can be performed for a specific purpose, as we did in the applications we presented.

The advantage of having an explicit structure to organise retrieved data for the user is that, even if hard-coded, it can be examined and changed to tailor the output of a search to different output devices by reasoning about which parts are the most important. For example, in both the tailored travel information application and in SciFly, non-essential information could be "hidden" in hyperlinks, if the device "real estate" did not allow for the presentation of all the text retrieved (or, in the case of paper brochures, which had to fit

in two pages, it could be suppressed (Paris et al. 2008a, b)). If flexibility in tailoring the output is needed, or to reason in other ways about the discourse, the organisational structure can be decomposed and encoded as plans, as presented in the NLG systems. A planning engine would then act as the orchestrator, with a search query being issued when appropriate. This, however, adds the cost of specifying the decomposition and having a planning engine.

Importantly, regardless of what underlying mechanism is used, when manually constructing these planning strategies or simply hard-coding structural relationships one should consider linguistic relationships, governing why information is juxtaposed with one another. This can inform decisions as to how to present the information.

In the NLG systems we presented earlier, as the designers of the systems had a priori knowledge of the underlying knowledge sources, we could ensure the result was coherent, relevant and looked like a human-authored text. To enable the application of such techniques to more open domains, where one may not have such a priori knowledge, we only need to relax the requirement on the output: when the underlying information sources are not known or do not have known structures, there is thus a risk of less than 100% recall and precision; hence there is also a risk of losing some coherence and relevance in the end result. This approach is illustrated in Fig. 8, and a hypothetical example of its use in to answer a complex information need was given in Fig. 7a.

To alleviate the burden of acquiring human-authored strategies, automatic approaches can be explored. Statistical methods have been used to uncover patterns in content structure—see for example Marcu (2000), Duboue and McKeown (2003), Barzilay and Lee (2004), Wan and Paris (2008), Sauper and Barzilay (2009). Duboue and McKeown (2003) demonstrated that a system is able to learn biographical schemas, given sufficient samples of biographical texts. More recently, Sauper and Barzilay (2009) wanted to produce Wikipedia pages automatically on selected topics—for example, descriptions of
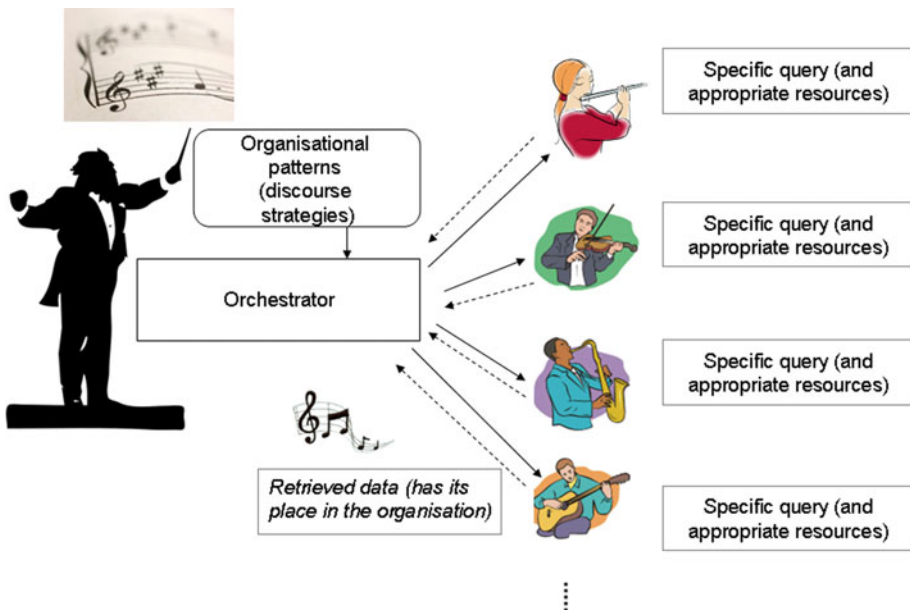

Fig. 8 The "orchestrator" approach to aggregating information

diseases. The availability of other such pages in Wikipedia provided a corpus that could be mined for structure. Having discovered the structure of such pages (e.g., high-level description, symptoms and signs, subtypes or causes, followed by prevention and diagnosis), the system could issue specific queries (e.g., "disease X and diagnosis") to web search engines to find appropriate text fragments (with sophisticated mechanisms for selection). This recent work demonstrates that this type of structural information can be automatically learnt.

Besides text analysis, further options for automatic acquisition include mining search engine query logs, click logs, user interface actions, and eye-gaze patterns to understand what users typically find interesting, what they read, and how the two relate to information needs and queries. Integrating any of these automatic models into standard NLG approaches, to reap their benefits, is still a topic for research.

6.2 Bottom-up approaches: finding data relationships in text

The previous examples in Sect. 5 were all top-down. For completeness, we describe ongoing research in bottom-up approaches, which analyse the information retrieved to determine relationships amongst information items. This is done in order to present the results in as coherent a manner as possible. Such work is strongly reminiscent of the corpus-based research on automatic acquisition of schemas described earlier. This can be schematically represented as in Fig. 9.

Some current aggregated search engines, including web search engines, do attempt to discover coherent organisations by clustering all or some of their results. Importantly, the notion of what to cluster upon is not fixed but left to a text analysis that is performed when the results are returned to the user. For example, in response to a query about a recent event, a system might retrieve all news items relevant to that event and cluster them by date (or date range). A sequential ordering amongst the cluster can then be inferred (based on the date), and the results could be presented in chronological order, thus giving a reader an understanding of the developments. Similarly, geographical results can be grouped and displayed on a map (Ringel et al. 2003; Alonso et al. 2007); or systems may cluster returned documents according to the topics they treat (Zeng et al. 2004). This would essentially provide an organisational principle to the underlying data, at a discourse (global) level.

Some work in summarisation can be seen as following this approach. For example, in their work on summarising a large document through extraction techniques, Wan et al. (2008) introduce the problem of sentence augmentation, in which a key sentence is *augmented* using information from neighbouring sentences. This integration, often requiring some text revision using paraphrasal operations, results in a newly generated summary sentence. This work demonstrates that supplementary information can be incorporated into
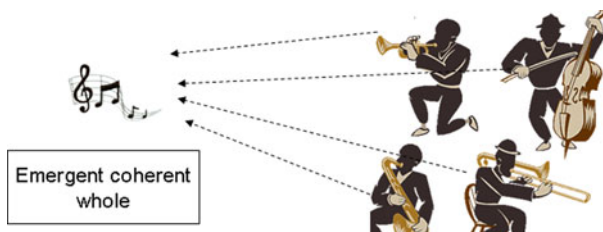

Fig. 9 Letting the structure emerge from the data as the basis for aggregation

a generated result, particularly at the level of a sentence. Ultimately, the summariser would perform this analysis at run-time on dynamically retrieved sentences in a bottom-up fashion. A hypothetical example of a bottom-up approach to answer a complex information need is given in Fig. 7b.

Similarly, in our work, we investigate how to automatically obtain supplementary information that can help a user understand a sentence (Wan and Paris 2008; Paris and Wan 2009). We term this elaborative summarisation. Built for web pages (currently Wikipedia), our system presents summaries of linked documents using the reading context of a user as a means to personalise the summary. In the domain of Wikipedia text, a summary of a linked document that is as yet unseen by the user is generated in order to find information that expands on the content of the linking sentence.

6.3 Exploiting both top-down and bottom-up: a hybrid approach

Each of the previous approaches has its strengths and weaknesses. In ongoing work, we explore further possibilities in combining top-down and bottom-up techniques in order to create coherent answer spaces from documents with both structured and unstructured data. Our current work looks at creating overviews about a topic, for someone new to a field or someone needing to catch up on recent developments. A top-down approach is first employed to direct a set of specific queries in potentially specific silos of information (e.g., organisational web pages; scientific articles; news). A bottom-up approach is then used to look for relationships amongst the items retrieved, and to produce summaries of topics of interest (such as individual organisations; people; or events).

Such an approach, which can accommodate a wide range of information but which still provides some structure, suits the company information task (Section 2) and other tasks requiring an overview of a large information space. It also suits tasks where there is unlikely to be the expertise needed to create detailed discourse plans, for example, in the enterprise domain. Our early experiments with this approach have accordingly focused on an enterprise task: that of enquiries staff at Australia's CSIRO. These staff must answer enquiries with reference to CSIRO's activities, those of other science organisations, recent events, and relevant people in the field; however, the information they need is scattered across a number of different repositories including some run by other organisations. A hybrid approach seems natural here.

In exploring hybrid systems that use both top-down and bottom-up methods, we are interested in answering a number of research questions including: how to employ bottom-up text analysis methods for different types of retrieved data (for example, web news articles and blog entries); how to acquire typical patterns of communications for novel applications (potentially requiring a metasearch function) for which examples of human-authored output are not available; and finally, how best to present aggregated results. We approach the first question through additional handling of different web-based text genres, and the latter two through the use of human-computer interaction techniques to understand better the information needs of the user, and thus the underlying communication goals of the aggregated result.

7 Conclusions

Current tools are capable of finding relevant documents, or passages, from a vast number of candidates but the onus of coordinating queries, extracting useful information from

documents returned, and noting relationships remains on the user. Although techniques such as metasearch and faceted search are some help, they do not address the problem of aggregating information into a coherent whole.

Natural language generation offers a different perspective on this problem. Research in this field has considered the problem of presenting information in a useful way, with reference to expected patterns of communication and theories of coherence. For example, by exploiting discourse strategies and planning techniques, a result can be built to answer an information need.

The two techniques—of information retrieval and natural language generation—take approaches at two ends of a continuum based on the genericity of the approaches, the appropriateness to task and context, and the understandability and naturalness of the results. Aggregated and focused search occupy something of a middle ground, and can benefit from both techniques. We have considered two possibilities: a "top-down" approach starts with a discourse plan, imposing an organisation on the answer space and specifying which information sources are to be queried and how. A "bottom-up" approach, by contrast, examines documents presented by a retrieval system and attempts to determine relationships in order to build a coherent answer. We have implemented aggregated search systems based on each of the top-down and bottom-up approaches, and are presently experimenting with a hybrid approach in systems that include both NLG and IR components.

We suggest that the two fields, of information retrieval and of natural language generation, can complement each other, to the benefit of each—and of search users, in particular in domains with complex information needs for which expected patterns of answers can be discovered.

References

Alonso, O., Baeza-Yates, R., & Gertz, M. (2007). Exploratory search using timelines. In *Proceedings of SIGCHI*. Workshop on Exploratory Search and HCI Workshop, San Jose, CA, USA, 28th April - 3 May (pp. 23–26). http://research.microsoft.com/en-us/um/people/ryenw/proceedings/ESI2007.pdf.

Austin, J. (1962). *How to do things with words*. London: Oxford University Press.

Barzilay, R., & Elhadad, M. (1997). Using lexical chains for text summarization. In *Proceedings of ACL/ EACL workshop on Intelligent Scalable Text Summarisation* (pp. 10–17).

Barzilay, R., & Lee, L. (2004). Catching the drift: Probabilistic content models, with applications to generation and summarization. In D. Marcu, S. Dumais, & S. Roukos (Eds.), *Proceedings of HLT-NAACL*. ACL—Association for Computational Linguistics (pp. 113–120). Boston, Massachusetts, USA.

Barzilay, R., & McKeown, K. R. (2005). Sentence fusion for multidocument news summarization. *Computational Linguistics, 31*(3), 297–328.

Bateman, J. A., Kamps, T., Kleinz, J., & Reichenberger, K. (2001). Constructive text, diagram and lay-out generation for information presentation. *Computational Linguistics, 27*(3), 409–449.

Callan, J. (2000). Distributed information retrieval. In B. W. Croft (Ed.), *Advances in information retrieval* (pp. 127–150). Dordrecht: Kluwer.

Choi, F. Y. Y. (2000). Advances in domain independent linear text segmentation. In *Proceedings of 1st meeting of the North American Chapter of the ACL* (pp. 26–33). Seattle, Washington, USA.

Colineau, N., Andrew, A. L., & Paris, C. (2004a). A task-sensitive user interfaces: Grounding information provision within the context of the user's activity. In *Proceedings of Advanced Visual Interfaces (AVI 2004)* (pp. 218–225). Gallipoli, Italy.

Colineau, N., & Paris, C. (2003). Task-driven information presentation. In *Proceedings of the Annual Conference of the Computer-Human Interaction Special Interest Group of the Human Factors and Ergonomics Society of Australia (OZCHI'03)* (pp. 138–146). Brisbane, Queensland, Australia.

Colineau, N., & Paris, C. (2007). Tailoring and the efficiency of information seeking. In *Proceedings of User Modelling* (pp. 440–444). Corfu, Greece.

Colineau, N., Paris, C., & Wan, S. (2001). Dynamic content presentation. In *Proceedings of Australasian Document Computing Symposium* (pp. 21–27). Coffs Harbour, NSW, Australia.

Colineau, N., Paris, C., & Wu, M. (2004b). Actionable information delivery. *Revue d'Intelligence Artificielle (RSTI - RIA), Special Issue on Tailored Information Delivery, 18*(4), 549–576.

Cutrell, E., Robbins, D., Dumais, S., & Sarin, R. (2006). Fast, flexible filtering with Phlat—Personal search and organisation made easy. In R. Grinter, T. Rodden, P. Aoki, E. Cutrell, R. Jeffries, & G. Olson (Eds.), *Proceedings of CHI* (pp. 261–270). Montréal, Canada.

De Carolis, B., de Rosis, F., Berry, D., & Michas, I. (1999). Evaluating plan-based hypermedia generation. In *Proceedings of European workshop on Natural Language Generation*. Toulouse, France.

Diaper, D., & Stanton, N. (Eds.). (2004). *The handbook of task analysis for human-computer interaction*. Lawrence Erlbaum Associates.

Duboue, P. A., & McKeown, K. R. (2003). Statistical acquisition of content selection rules for natural language generation. In *Proceedings of Empirical Methods in Natural Language Processing* (pp. 121–128). Sapporo, Japan.

Green, N., Carenini, G., Kerpedjiev, S., Roth, S., & Moore, J. (1998). A media-independent content language for integrated text and graphics generation. In *Proceedings of Workshop on Content Visualisation and Intermedia Representations* (pp. 69–75). Quebec, Canada.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts, Vol. 3*. New York: Academic Press.

Hearst, M. (2006). Design recommendations for hierarchical faceted search interfaces. In *ACM SIGIR Workshop on Faceted Search*. Seattle, Washington, USA.

Isard, A., Oberlander, J., Androutsopoulos, I., & Matheson, C. (2003). Speaking the users' languages. *IEEE Intelligent Systems Magazine, Special Issue "Advances in Natural Language Processing", 18*(1), 40–45.

Jones, R., Rey, B., Madani, O., & Greiner, W. (2006). Generating query substitutions. In *Proceedings of WWW*, ACM, New York, NY, USA, pp. 387–396. doi: http://doi.acm.org/10.1145/1135777.1135835.

Kamps, J., Geva, S., Trotman, A., Woodley, A., & Koolen, M. (2008). Overview of the INEX 2008 ad hoc track. In S. Geva, J. Kamps, & A. Trotman (Eds.), *Proceedings of INEX* (pp. 1–28). Dagstuhl, Germany.

Lalmas, M., & Murdock, V. (2008). *SIGIR workshop on aggregated search*, Singapore. http://www.yr-bcn.es/sigir08/.

Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text, 8*(3), 243–281.

Marcu, D. (2000). *The theory and practice of discourse parsing and summarisation*. Cambridge: The MIT Press.

McKeown, K., & Radev, D. (1995). Generating summaries of multiple news articles. In *Proceedings of SIGIR*, ACM Press (pp. 74–82). Seattle, Washington, USA.

McKeown, K. R. (1985). *Text generation: Using discourse strategies and focus constraints to generate natural language text*. Cambridge: Cambridge University Press.

Mohammad, S., Dorr, B., Egan, M., Hassan, A., Muthukrishnan, P., Qaazvinian, V. et al. (2009). Using citations to generate surveys of scientific paradigms. In *Proceedings of NAACL/HLT, Association for Computational Linguistics* (pp. 584–592). Boulder, Colorado, USA.

Moore, J., & Paris, C. (1993). Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Journal of Computational Linguistics, 19*(4), 651–694.

Moore, J., & Swartout, W. (1991). A reactive approach to explanation: Taking the user's feedback into account. In C. Paris, W. Swartout, & W. Mann (Eds.), *Natural-language Generation in Artificial Intelligence and Computational Linguistics* (pp. 3–48). Dordrecht: Kluwer.

Paris, C. (2006). *Search and delivery in context: Another perspective on CL and IR*. Invited talk at Coling/ACL workshop on Computational Linguistics and Information Retrieval, Sydney, Australia.

Paris, C., & Colineau, N. (2006). *Scifly: Tailored corporate brochures on demand*. CSIRO ICT Centre Technical Report TR 06/26.

Paris, C., Colineau, N., Lampert, A., & Giralt Duran J (2008a). Fit it in but say it well!. In *Proceedings of Annual Workshop of the Australasian Language Technology (ALTA)* (pp. 133–141). Hobart, Tasmania, Australia.

Paris, C., Colineau, N., Lampert, A., & Giralt Duran J. (2008b). Generation under space constraints. In *Proceedings of COLING (Computational Linguistics)* (pp. 127–129). Manchester, UK.

Paris, C., Colineau, N., Lampert, A., & Vander Linden, K. (2009). Discourse planning for information composition and delivery: A reusable platform. *International Journal of Natural Language Engineering, 16*(1), 61–98.

Paris, C., Lampert, A., Lu, S., & Wu, M. (2005). Enhancing dynamic knowledge management services—Tailored documents. Tech. Rep. Commercial-in-Confidence 05/034, CSIRO ICT Centre.

Paris, C., & Wan, S. (2009). Capturing the user's reading context for tailoring summaries. In *Proceedings of International Conference on User Modelling*. Adaptation and Presentation (UMAP) (pp. 337–342). Trento, Italy.

Paris, C., Wan, S., Wilkinson, R., & Wu, M. (2001). Generating personal travel guides—And who wants them? In M. Bauer, P. Gmytrasiewicz, & J. Vassileva (Eds.), *Proceedings of User Modelling* (pp. 251–253). Sonthofen, Germany.

Paris, C., Wu, M., Vercoustre, A., Wan, S., Wilkins, P., & Wilkinson, W. (2003). An empirical study of the effect of coherent and tailored document delivery as an interface to organizational websites. In *Proceedings of Apdaptive Hypermedia Workshop* (pp. 133–144). Pittsburgh, Pennsylvania, USA.

Paris, C. L. (1987). Tailoring descriptions to a user's level of expertise. *Journal of Computational Linguistics, 14*(3), 64–78, http://www.aclweb.org/anthology/J/J88/J88-3006.pdfM.

Qazvinian, V., & Radev, D. (2008). Scientific paper summarization using citation summary networks. In *Proceedings of COLING (Computational Linguistics)*. Manchester, UK.

Radev, D., Otterbacher, J., & Qi, D. T. H. (2003). Mead reducs: Michigan at DUC 2003. In *Proceedings of Workshop on Summarization, Document Understanding Conference*.

Ringel, M., Cutrell, E., Dumais, S. T., & Horvitz, E. (2003). Milestones in time: The value of landmarks in retrieving information from personal stores. In *Proceedings of Interact*.

Sacerdoti, E. (1977). *A structure for plans and behavior*. New York: American Elsevier.

Sauper, C., & Barzilay, R. (2009). Automatically generating wikipedia articles: A structure-aware approach. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP, Association for Computational Linguistics* (pp. 208–216). Suntec, Singapore, http://www.aclweb.org/anthology/P/P09/P09-1024.

Scott, D., & de Souza, C. (1990). Getting the message across in RST-based text generation. In R. Dale, C. Mellish, M. Zock (Ed.), *Current research in NLG* (pp. 119–128). London: Academic Press.

Shaw, J. (1998). Clause aggregation using linguistic knowledge. In *Proceedings of the9th International Workshop on Natural Language Generation*, Niagara-on-the-lake,Ontario, Canada, August 5–7. pp. 138–147.

Shushmita, S., Lalmas, M., & Tombros, A. (2008). Using digest pages to increase user result space: Preliminary designs. In *Proceedings of SIGIR 2008 Workshop on Aggregated Search*, Singapore.

Szlavik, Z., Tombros, A., & Lalmas, M. (2007), Content and structure summarisation of XML documents for effective information access. In *Proceedings of DELOS conference on Digital Libraries*, Pisa, Italy.

Taboada, M., & Mann, W. (2006). Applications of rhetorical structure theory. *Discourse Studies, 8*(4), 567–588.

Teevan, J., Alvarado, C., Ackerman, M. S., & Karger, D. R. (2004). The perfect search engine is not enough: A study of orienteering behaviour in directed search. In *Proceedings of Conference on Human Factors in Computing Systems* (pp. 415–422) Vienna.

Teufel, S., & Moens, M. (2002). Summarizing scientific articles: Experiments with relevance and rhetorical status. *Computational Linguistics, 28*, 409–445.

Thomas, P. (2008). *Server characterisation and selection for personal metasearch*. PhD thesis, Australian National University.

Trotman, A., Geva, S., & Kamps, J. (2008). In *Proceedings of SIGIR workshop on focused retrieval: Question answering, passage retrieval, element retrieval*, Singapore.

Wahlster, W., André, E., Finkler, W., Profitlich, H., & Rist, T. (1993). Plan-based integration of natural language and graphics generation. *Artificial Intelligence, 63*(1–2), 387–428.

Wan, S., & Paris, C. (2008). In-browser summarisation: Generating elaborative summaries biased towards the reading context. In *Proceedings of Association for Computational Linguistics: Human Language Technologies*. Short papers (pp. 129–132). Columbus, Ohio, USA. June 2008.

Wan, S., Paris, C., & Dale, R. (2009). Whetting the appetite of scientists: Producing summaries tailored to the citation context. In *Proceedings of Joint Conference on Digital Libraries* (pp. 59–69). Austin, Texas, USA.

Wan, S., Dale, R., Dras, M., & Paris, C. (2008). Seed and grow: Augmenting statistically generated summary sentences using schematic word patterns. In *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing* (pp. 543–552). Honolulu, HI: Association for Computational Linguistics.

Wilkinson, R., Wu, M., & Fuller, M. (2001). Search performance in question answering. In *Proceedings of SIGIR*. New Orleans, LA: ACM.

Xu, J., & Croft, W. B. (1996). Query expansion using local and global document analysis. In *Proceedings of SIGIR* (pp. 4–11). Zurich, Switzerland: ACM.

Zaragoza, H. (2008). Yahoo! reference: Rich search results for informational queries. Invited talk at the SIGIR Workshop on Aggregated Search, Singapore.

Zeng, H. J., He, Q. C., Chen, Z., Ma, W. Y., & Ma, J. (2004). Learning to cluster web search results. In *Proceedings of SIGIR*, ACM, New York, NY, USA, pp. 210–217, doi: http://doi.acm.org/10.1145/1008992.1009030.