Chapter 6

Context and HyperContext

6.1 Introduction

Over the last few years, the role of *context* in Artificial Intelligence (AI) domains has received increasingly important status. It is generally agreed that context (for example, the context of a discourse) has always been implicitly represented in computer systems. Until recently, however, context has not been explicitly represented and manipulated in order to achieve a better understanding of an interaction between a human and a computer, or indeed, between two or more interacting computer systems.

Consider the simple statement, “The King of France wears a wig”. The statement is meaningful (as long as the reader is familiar with English), but its truth (in terms of facts about the world) is dependent on the time of the utterance, location, and other factors which may influence its intended meaning[1].

Context, however, is not simply about grounding statements to temporal events. In order for participants in a discourse to properly understand each other (where the listener understands the speaker in the way the speaker intends), each participant usually constructs a (partial) model of the other. The model may contain assumptions (about the other’s intentions, knowledge directly relevant to the topic of discourse, referents, and so on). The model may also contain facts about the other participant as more of the assumptions are confirmed, and as new information is derived from what is known and assumed. Communication breakdown (such as a misunderstanding) can occur when the speaker and listener have constructed incorrect models of each other. Sometimes, the

[1] It could be, for instance, that in certain societies, wig-wearing is a reference to an attempt to cover up the truth, so the intended meaning of the utterance is not merely a statement of fact, but is instead a comment about the trustworthiness of the person concerned.

assumptions may go uncorrected (in which case there may be confusion, an argument, or one of the participants may feel insulted or hurt), whereas on other occasions, the mistake will be noticed (perhaps as one of the participants is making utterances which contradict the other’s assumptions, and clarification is sought), in which case the models can be re-aligned.

John McCarthy’s seminal work on formalising context [63] has motivated much research in logics of context and the use of context in problem solving.

In HyperContext, we require context to support the distinction between different interpretations of the same document to facilitate the construction and maintenance of a user model which represents the user's short-term interests. From this user model, we derive a query which represents information the user wants to locate. Context will enable the browsing environment to guide the user to relevant information by recommending links to follow, or to provide context-free "See Also" links to information which exists in a different context but which the user may consider relevant.

After discussing context in the literature, we discuss the approach to context taken in HyperContext and describe the composition, dependencies, and effects of context.

6.2 Background

McCarthy is accredited with stimulating research in the area of context in Artificial Intelligence (AI). He thought that reasoning in context would make AI systems less brittle, so they would be more able to cope in dynamic environments; the scope of terms could be limited to contexts, in order to facilitate interaction between software components that cannot easily communicate because of inconsistencies in the different terminologies, or ontologies, used by each system; and that AI systems would be able to transcend the limits of their knowledge.

The major research to arise from this initial work is Cyc [44], a knowledge-based representation of common-sense information, and the definition of a logic for reasoning with it. Other research falls predominantly into one of two categories. Either the logic proposed by McCarthy and implemented by Guha is extended, or else the notion of context is applied to different AI domains, such as Natural Language Processing, to solve previously difficult-to-solve problems.

At the simplest level of abstraction, a context partitions a dataset so that statements can be made and understood within specific contexts without requiring the use of *universal* *statements*[2]. For example, the statement “He wears a wig” is ambiguous if intended as a universal statement. Additional information may be progressively added to disambiguate the statement, such as “The King wears a wig”, “The King of France wears a wig”, and “Louis XIV, the King of France wears a wig”. Each statement is “more universal” than the previous, because less assumptions need to be made by the reader to understand what the writer intends. However, if it is required to make all assumptions explicit within a statement, statements would be extremely cumbersome and conversation would be an extremely tedious pastime! In a context in which it is assumed that the topic of discussion is, in fact, Louis XIV, the King of France from 1643 to 1715, then in the statement "He wears a wig" it can be assumed that the anaphoric reference is to Louis XIV, and therefore, it does not need to be made explicit in the statement.

Within a context, assumptions that are needed to correctly interpret statements are made explicit, and are part of the statements' environment. Consequently, it is possible for identical statements in different contexts to have different interpretations.

6.2.1 What is “context”?

“Context” defies a universal definition. McCarthy has 'rejected the idea of defining what context is' [64], although he describes what contexts do. In other areas, it is accepted that context can only be spoken of in reference to its use [12], and the various definitions of context are specifically grounded in the domains within which they are used.

> Context is something surrounding an item and giving meaning to this item... context acts then on the relationships between items than on items themselves. [12]

> ... we will accept a very general notion of context as a collection of "things" (parameters, assumptions, presuppositions, ...) a representation depends on. [40]

Brézillon describes Edmondson and Meech’s view of context as 'Context is what gives meaning to data and "contextualization" is the process of interpreting data, transforming data into information' [11]. A logic of context makes the “something” called context explicit, and defines the relationship between the context and the data it surrounds, as well as the operations on contexts and data interpreted in context, which convert it into information.

---

[2] A statement which is unconditionally true.

In Natural Language Processing, context can be viewed as 'a beliefs environment, a structure of nested belief-spaces for supporting the interpretation and production of natural language utterances (and other actions)' [34]. Giunchiglia [38], similarly, but for the general area of goal-based reasoning, states that 'a *context c* [is] *that* **subset of the complete state** *of an* **individual** *that is used for reasoning about a given goal*'. However, contexts are rich, in that we cannot possibly know all, much less completely encode, the factors that are required to completely describe the context surrounding a conversation, for instance. Although attempts to describe context are incomplete, there is universal agreement that 'the explicit use of context limits the domain of validity of the acquired knowledge and indicates the correct moment of use' [11].

Research into the substance of context falls into two broad categories. Inspired mainly by John McCarthy, the first major research area is concerned with extending existing logics or formulating new logics to reason with information-in-context. The second category deals with the explicit representation of context in applications to improve reasoning in specific domains.

Giunchiglia [39] provides a good introduction to contextual reasoning in which two major roles for context are identified, pragmatic and cognitive contexts, although the authors argue that the role of pragmatic context is subsumed by cognitive context. Pragmatic context, identified largely from the theories of Bar-Hillel and Kaplan (cited in [39]), deals with context as a part of the state of the real world, whereas cognitive contexts (McCarthy, Sperber and Wilson, Kokinov, Fauconnier, and Dinsmore, all cited in [39]) treat context as part of agent's representation of the world.

6.2.2 A logic of contexts

Following [63], central to a logic of contexts is that a proposition (a sentence, a statement) *p* is asserted to be true in a context, which is itself asserted in some outer context. The context contains all the assumptions that are needed to make the proposition both true and understood as intended.

Apart from statements which are true in context, there are also terms whose *value* is context dependent. For example, an optimist might consider a glass to be half full, whereas a pessimist might consider the same glass to be half empty. As this simple example demonstrates, the ability to contextualise statements allows a data space to maintain inconsistent, and even seemingly contradictory statements and observations about the universe.

Predicates are true in a context, and contexts are themselves asserted within outer contexts. The regression is infinite, but McCarthy shows that this is harmless [65]. However, for most implementations of context, assuming an outermost context is reasonable, although this effectively means that the model of the world represented is closed.

In a system that manipulates context, when the system is in an outer context, it may make assumptions about sentences that are true and give assumed values to terms (in the sense of Default Logic [83]). When the assumptions need to be asserted, perhaps because reasoning cannot continue, then the appropriate context can be *entered*. Once a logical conclusion q has been derived, then the context can be left and q can be asserted in the outer context.

Although partitioning datasets using contexts is useful because the same sentences in different contexts can be properly understood as intended, this in itself is not particularly powerful - we would like to be able to relate sentences and terms across contexts, as in this way a system may be able to increase its reasoning power. *Lifting* relations permit such referencing.

A lifting relation is a rule which can be used to relate inconsistent or otherwise incompatible definitions of the same thing. For example, consider a context in which Da Vinci's *The Last Supper* [26] is described, using its physical dimensions. The location of the painting is unimportant, although it can be assumed to have one. However, in a context in which the painting has been stolen, the last known location of the painting is important. In order to relate the two descriptions of the painting, it is necessary to be able to lift the description of *The Last Supper* into a context in which the location is made explicit. Many lifting relations are used to *situate* objects described in a context. Most of the time, people are able to discuss concepts, without reference to concrete examples. It is possible for us to talk about a meal, a painting, or a telephone call, where the situation details are omitted, although thanks to our common sense knowledge we know a meal is eaten off a plate with cutlery, a painting has a home, a telephone call means that a successful connection has been made. For instance, if I claim to have spoken to Mike on the phone, others would expect that Mike had also spoken to me, and it would be a surprise if he denied that the call had been made, or claimed to be nowhere near a phone when the call was made. However, there are occasions when the context in which a discourse takes place requires that the assumptions are made explicit, and lifting rules permit this.

Sentences that a system unconditionally believes to be true are located in the outermost context. Although, according to McCarthy, there is no tangible outermost context, it is useful to consider systems as being in an implicit outermost context. However, the implicit existence of an outermost context is not limiting. The system can always *transcend* the outermost context by creating a new context in which the assumptions of the old context are relaxed or changed. With the ability to transcend context, systems are able to relax or change universal assumptions and statements.

6.2.3 The applications of context

The benefits of context are far-reaching, and all the possibilities have not yet been considered. Making context explicit and manipulating it is useful in Natural Language Processing and discourse analysis, to support not only anaphora resolution, but also to understand idiomatic expressions and limit the meaning of terms of reference ([34] and [73]). Context has a role to play in Constraint Satisfaction Problems [92], to describe solutions that can change depending on the environment of the variables. Case-Based Reasoning (CBR) is another area in which context is already implicitly present, and which can benefit from making context explicit. In CBR, one of the challenges is to identify existing cases which resemble the description of the current problem in order to apply previous solutions to solve the current problem. Cases are indexed by key features, which are expected to make the case “stand out” and be easily identifiable when looking for similarities between them and the current problem. Schank’s theory of Dynamic Memory [80] concerned TOPs (Thematic Organization Packets). In CBR, a TOPs approach would enable a case which appears to bear little resemblance to the current problem to participate in its solution if the solution to the first problem could be used to solve the current problem. Context can be used to determine which key features of a case to index so that it will be found whenever it is relevant, but will be ignored when it is irrelevant [51].

Co-operating software agents (Intelligent Agents) which use different ontologies to describe their worlds (even though they may operate in overlapping worlds, or in different worlds which require similar skills in which to operate) would be able to contextualise their knowledge and share it in order to facilitate the co-operation. Lifting relations could provide mechanisms for co-operating agents to be able to understand each other [82].

Other areas in which context can be modelled and used are User Modelling in large, heterogeneous domains, where what is learnt about a user can be limited to relevant contexts, yet at the same time, a system can use lifting rules to apply what is known about the user in one area to assumptions about what she might be interested in in another, possibly unrelated area. Information Retrieval systems take a user query and apply it to a representation of a document base to determine which documents are likely to be relevant to the query, and hence the user's interests. In Information Retrieval (IR) different users with different requirements may express their information need using identical queries. IR systems would normally return the same set of documents to each user. If the context of the query can be modelled and matched against the contexts in which interpretations of the documents are relevant, then different users with different requirements and expectations who use the same query may well be presented with different sets of relevant documents.

As context is used to partition a dataset, it can be used in knowledge based systems, as in Cyc [44], relational databases, hypertext hyperspaces, machine learning, decision support, data mining, game theory, human-computer interaction, intelligent tutoring systems, and digital libraries, to name but a few application areas.

6.3 Context in HyperContext

In HyperContext we want to provide users with an environment that actively participates in and supports the user's search for relevant information. We use *context* to guide the user to relevant information; to determine what information is relevant; and, to automatically estimate the user's current interest.

We have liberally referred to "context" earlier in this thesis without really explaining it. We know that context enables HyperContext to support multiple representations of the same document as interpretations. We know that an interpretation, created by a member of the HyperContext community, contains the information required to describe a document in context. We know that the same document can be linked to different children in different contexts, so that a link to an expert treatise on a topic in one context can be replaced with a link to a gentler introduction to the same topic in another context. In the remainder of this chapter, we will describe what context means in HyperContext, and the rôle it plays.

6.3.1 Interpreting information in context

A document is the smallest unit of information which can be interpreted in HyperContext. Whenever a browser accesses a document during a HyperContext session, we say that the document is interpreted. The interpretation is represented by a vector of term weights, which describes the document in context, and a set of out-links which are $node_1$-label-$node_2$ triples, where $node_1$ is the name of the document under interpretation, label is a non-zero weighted term in the interpretation which acts as the link anchor, and $node_2$ is the name of the destination document. Finally, we have defined context to be a node-label pair (Chapter 4.3).

These structures allow us to talk about interpreting a document in context; partitioning the hyperspace; contextual relevance; and guiding a user to contextually relevant information. However, the feeling persists that we should be able to do something more with context. A HyperContext hypertext is constructed by its community of users. HyperContext does not create links between nodes - its users do. HyperContext does not provide the interpretation for a document in context - a user does. Why does HyperContext not automatically provide an interpretation for document, given a context? Why does HyperContext not extend a user's browsing space by automatically linking to documents which have an interpretation relevant to the context, so that a user can follow a dynamic link?

Dynamic links

Let us assume that a link can have an unnamed destination, so that at run-time the most appropriate destination for the link can be evaluated. Traversing such a link would result in the user being taken to the most relevant interpretation of a document. The process of creating a link normally involves the selection of a label to act as the link source from an interpreted document, a destination document, and a vector of weighted labels which describe the interpretation of the destination document. Assume that creating a link with an unidentified destination involves the same steps just described, except that the identity of the destination document is not retained. The link might be represented as the triple $node_1$-label-EVAL, where EVAL is a run-time instruction to search for, or create, a relevant interpretation of a document using the vector specified by the creator of the link. In HyperContext we represent interpretations and a user query using vectors of term weights. It is therefore conceivable to consider that in a given context a document's interpretation is equivalent to the query for which the user had selected the document as being the most relevant.

Let us assume that upon encountering a $node_1$-label-EVAL link, an information retrieval system will search through existing interpretations of all documents to locate the one which is most relevant. Although the solution appears attractive, we consider that it is not suitable. The located existing interpretation was created in a different context, and so, given that in HyperContext context plays a part in determining relevance, we would not create a context-sensitive link to that interpretation, although HyperContext can recommend it via a context-free "See Also" link. In Information Retrieval-in-Context and

Adaptive Information Discovery, an interpretation which appears to be relevant to the user query, but does not exist on the same context path as the node the user is currently visiting, is recommended to the user only as a context-free "See Also" link, and it is never recommended through a context-sensitive embedded link.

Another problem associated with offering dynamic link destination evaluation is manifested if *all* links are dynamic. An interpretation of a document is normally created by the user who decides that the document should be the destination of a link. However, if all links have no specified destination, then no document will have pre-existing interpretations. If interpretations do not exist, then an information retrieval system cannot even locate superficially relevant interpretations as a result of evaluating a link destination. This implies that the interpretation of a link destination would need to be created dynamically by HyperContext. This is roughly equivalent to requiring an arbitrary document to be interpreted in an arbitrary context. The following subsection describes how an arbitrary document might be interpreted in an arbitrary context.

Arbitrary interpretations of documents

Consider a HyperContext hypertext where all links dynamically bind to their ideal destination on demand. Whenever a user accesses a document, a dynamic link is created when the user selects an arbitrary region of the document to act as a link source anchor. The accessed node and the label (which represents the region selected as the link source anchor) form the context in which a destination document is to be interpreted. HyperContext must evaluate a user's interest so that it can automatically locate a document to act as the link destination. Interpretations of documents will not pre-exist, so whenever a document is accessed it must be interpreted so that HyperContext will be able to estimate a user's short-term interest. A similar approach is taken in Chapter 8, which describes the automatic conversion of a Web site to a HyperContext site for evaluation purposes. For the time being, we will assume that it is possible to automatically interpret a document given a context, although an automatic interpretation of a document is derived objectively, whereas an interpretation created by a user will be subjective.

At the start of a new context session, the user model representing the user's short-term interest will be empty. When the first document is accessed, it will normally be accessed in the context bottom. Assume that the interpretation derived for the document in the context bottom is the equivalent of the term weight vector derived in a standard vector space model of information indexing, using the TFxIDF metric described in Chapter 5.2. The user model is instantiated to this vector. When the user selects an arbitrary region of the document to act as the link source anchor, the user model is modified to give greater weighting to the label associated with the region, as are the weights of terms which occur in close proximity to the anchor. To evaluate the link destination, a query is abstracted from the user model, and the document base is searched. However, HyperContext requires that the user is taken to the most relevant *interpretation*. This implies that either the search is conducted to locate the most appropriate *document* which is then interpreted, or else each document is first interpreted and the most relevant interpretation is returned. The first option, to locate a relevant document and to then interpret it, appears to be contradictory to the whole idea of HyperContext, so we investigate the alternative.

When a user creates an interpretation of a document, she is free to use any terms from the vocabulary, even if they do not explicitly occur in the document. For the purposes of the current argument, we will impose a restriction on the terms used to describe an interpretation, by stating that the terms used must occur in the document. Normally, an interpretation is constructed by selecting those regions of a document which the users consider to be pertinent to their information foraging task. The selected regions are then indexed and represented by a vector of term weights. Although the user can normally modify term weights, we will also remove this ability so that HyperContext can automatically create interpretations of documents. If an interpretation is a representation of selected regions of a text document, then all possible interpretations of the document can be automatically derived by dividing it into all its possible regions (the individual words), and then systematically combining all the regions, each unique combination forming an interpretation. Interpretations will range from single words to the most expansive interpretation which is the entire document (in the current argument, this is the interpretation of the document in the context bottom). The task is then to find the most appropriate interpretation of any document to form the destination of the link.

At the start of a new context session, a user will access a document in the context bottom. The interpretation of this document will be the most expansive interpretation. The short-term user model is initialised to this interpretation. The user selects a region to act a dynamic link source anchor, and the user model is updated to reflect the user's interest in the terms which occur in the selected region. A query is generated from the user model and an exhaustive search of all possible interpretations reveals the one which is most relevant. The user then accesses this document and browsing continues in the same way. Whereas this might appear to be the ultimate adaptive hypertext system, we argue that for HyperContext this method is not desirable, for reasons outlined below.

In an environment where all possible interpretations of a document are derived automatically from the contents of the document, rather than by manually deriving interpretations based on contexts in which the document naturally exists, there is no information available to HyperContext for it to decide on the likelihood of an interpretation of a document actually being "reasonable". When an interpretation is created manually, the system (and its users) can assume that the created interpretation is "reasonable", although nothing can be said about those possible interpretations which do not exist (they might not exist either because they are "unreasonable" or because nobody has created them yet). In a HyperContext environment in which users participate in the creation of interpretations, the processes of creating a context, a link connecting a parent document to a child, and the creation of the interpretation occur simultaneously. The vector which represents the interpretation of the document in context can be equated to the query the user "would have asked", if only they had known what to ask, because it is convenient for us to reason about it in this way. However, we are imposing no requirement for a correspondence between the user model representing the user's short-term interests and the description given to the interpretation of the document - the user is free to describe the interpretation in any way, so long as it is a reasonable representation of what she considers the essence of the document to be in this context. However, when we consider the proposed way of automatically locating the most appropriate interpretation as the destination of a dynamic link, there is necessarily a requirement for a high correspondence between the user model and the interpretation. This implies that instead of being able to learn what the user may be interested in by observing differences between the user model and the interpretation the user elects to visit, the automatically selected interpretation merely confirms what the user model already contains.

Favouring predetermined links and interpretations

We have briefly argued against links and links destinations being dynamic - the main reasons being the importance of the human users' participation in the creation and maintenance of a HyperContext hypertext. Having said that, there is scope in HyperContext to provide dynamic links (as context-free "See Also" links) mainly through Adaptive Information Discovery (AID). The search mechanisms provide the main way through which a context path can be extended or created, but the user decides which document found through search is subjectively relevant and the way in which it should be interpreted.

6.4 The nature of context in HyperContext

We will subject the notion of "context" to closer inspection, predominantly to tackle issues such as what context is, what a context contains, how a context influences the interpretation of a document, and what dependencies contexts have.

At face value, a context is just the name of an environment within which a document can be interpreted. Although the context name can be arbitrary, the convention we use is that the name of a context is derived from the combination of the name of a label which acts as a link anchor in the source document, and the name of the source document. The set of possible contexts is $(\Delta \times \Gamma) \cup \{\perp\}$, where Δ is the set of unique document names (obtained from the Object Layer), Γ is the set of unique label names (the vocabulary of terms), and $\{\perp\}$ is a one member set containing the context bottom. Likewise, the set of possible links is $\Delta \times \Gamma \times \Delta$.

The naming convention we use implies that a context does not exist independently of a link. We have also established that an interpretation does not exist independently of a link, either. There is a direct dependency between a link, an interpretation, and a context, so much so that it is not possible for one or more of them to exist without all of them existing.

In Chapter 4.3 we represent a link by the triple node$_1$-label-node$_2$, and context is represented by the pair node-label. A link implicitly contains the context of the destination, which is made explicit by extracting it from the link. Traversing a link causes the destination document to be interpreted, which is a function of getInterpretation(node$_2$, context(node$_1$, label)).

In different interpretations of the parent (node$_1$, in this case), label may be free (not linked); linked to a different destination; linked to the same destination; or even zero-weighted. Whether node$_1$-label forms a valid context, and how it behaves if it does, is dependent on the interpretation of node$_1$. In addition, a *document* is interpreted as a result of a link traversal. A more accurate representation of a link, therefore, is the triple node-label-document, where document is the name of the destination document in its uninterpreted form. The act of traversing the link causes the interpretation of document in the context of node-label. To ensure the inseparability of a context and a link, we say that the context exists on the link, and that the context contains the assumptions (in the form of a vector of term weights, and a set of out-links) which enable the destination document to be interpreted. An inevitable implication of this organisation is that contexts are nested but, in keeping with [65], this is harmless, and we do not need to fully expand a document's context in order to "understand" the interpretation.

A context, a link, and the interpretation of the destination document appear to be inseparable from the interpretation of the document which contains the link source and which, consequently, indirectly provides the context for the interpretation of the destination document. For *practical* purposes, they are inseparable. However, in the definition of a link (node-label-document), node does *not* refer to a specific interpretation of a document. Rather it refers to *any* interpretation of the document named node. Consequently, we cannot unambiguously tell in which context node itself was interpreted from a link node-label-document in isolation. There may be more than one possible context of node in which node-label-document is an out-link, and the link gives insufficient information for us to determine the exact context in which node was interpreted. This means that although several interpretations of the document named node may have node-label-document as a link, only one interpretation of document in the context node-label is possible. Although this is seemingly a disadvantage, stating "it is not possible for the same document to have different interpretations in the same context" simply highlights the implausibility of somehow allowing the same document to have different interpretations in the same context. It can be argued that node-label does not, in fact, refer to a unique context, but gives the same name to a potentially infinite number of different contexts where each context named node-label exists in a different stratum of super-contexts, and so it should be possible for "different" instances of a node-label context to interpret the same document in different ways. Although we agree with this conclusion, we object to its implications on the grounds that in order for the same context node-label in different super-contexts to support different interpretations of the same document, each context would need to be referred to by its fully qualified, or otherwise uniquely identifiable, name. This would be tantamount to using universal statements which in Section 6.2 we discounted as being impractical. Therefore, although it may appear to be a disadvantage, a given node-label-document link can support only one interpretation of a given document.

To conclude this section, a link is a node-label-document triple, from which node-label is extracted to form an explicit context within which to interpret a document. The context contains a vector of term weights which are used to describe the interpretation of document when the link is traversed, and a set of out-links for the interpreted document. The link and the interpretation of the destination document, and therefore the context, are specified by a member of the HyperContext community.

6.5 Relevance and context

The Traditional Information Retrieval (TIR), Information Retrieval-in-Context (IRC), and Adaptive Information Discovery (AID) search mechanisms all take a query representing a user interest and search for, and return to the user, relevant information. Chapter 5 has described these search methods, mainly from the perspective of the user. In this section we will briefly revisit the search mechanisms and explore them from the perspective of information in context.

A context path is a path that can be followed by a browser by consecutively following a link from an interpreted document to another until an interpreted document without out-links is reached. The context path had been constructed by previous HyperContext users by extending new links from existing interpreted documents to other documents for which they provided a new interpretation. Whenever a new interpretation of a document is created, the user describes the interpretation using terms that are pertinent to her. If a new complete context path is created, the first document in the path will have no parent, so the document is described in the HyperContext's outermost context, bottom.

Nodes on a context path are relevant to each other, not necessarily because a similarity measure would find a high degree of similarity between them, but because a human user has implied relevance between a parent and a child by creating a link between them. However, when a user creates an interpretation, she does not have to consider how to relate the new interpretation to nodes earlier in the context path, so the relevance of two nodes diminishes as a factor of the *contextual distance* between them. There are two types of relevance. The first, *contextual relevance*, describes nodes which are relevant to a user query and are contextually close to the user's current location in hyperspace. The second, *superficial relevance*, describes nodes which are relevant to the query, but are contextually distant from the user. The contextual distance between two nodes is greatest if the only way of reaching one document from another is through the context bottom.

We implicitly take contextual distance into account when we construct a model of the user's short-term interests in Chapter 5.8.2, and as further discussed in Section 6.7.2. The influence of nodes accessed earlier in a context session have a diminished influence in the user model, which is represented by the weight given in the scale of confidence in the interpretation of a given node actually contributing accurate information to the representation of the user's interests. However, the HyperContext framework does not yet take full advantage of contextual distance between nodes. We consider a node to be contextually relevant only if it is relevant to a user query and it is on the same context path as the node the user is currently visiting (subject to a maximum user definable number of link traversals). Superficial relevance relates to nodes which occur on different context paths or which exist on the same context path but which would exceed the maximum number of link traversals to reach. A node is contextually non-relevant if it is contextually distant and non-relevant to the query. A context switch occurs if a user hyperleaps from one node to another and the destination node is non-relevant or only superficially relevant. The HyperContext framework needs to be extended to define a node which is contextually close to the user, but which is non-relevant to the user query.

6.6 Guiding users to relevant information

HyperContext distinguishes between information that is contextually relevant and information that is superficially relevant (Section 6.5). AID can recommend superficially relevant information via "See Also" links, but the user can be guided to contextually relevant information through link and path recommendation.

Following a search for information it is necessary to determine whether the interpretation to be recommended is contextually or superficially relevant. To do so, we expand the context of the relevant interpretation. Information Retrieval-in-Context (IRC) and Adaptive Information Discovery (AID) are complementary search methods. In each case, the contexts of the two interpretations (one from which the search is activated, the other which is relevant) are compared to establish if they are connected via a context path. In IRC, the information of prime interest is that which is contextually relevant, so the search for relevant information can be accomplished by following all context paths starting from the search node until the first relevant interpretation is located. Once a path to relevant information is found, the link to follow in each node on the path is highlighted. Alternatively, all contextually relevant interpretations can be displayed to the user who can then choose to follow a recommended context path to one or more of them.

In AID, we are also interested in superficially relevant information which can be recommended via "See Also" links. Technically, superficially relevant interpretations are the set of interpretations relevant to the user query which are not contextually relevant. AID is a combination of TIR and IRC solutions. TIR will locate all relevant interpretations (as it does not distinguish between contextual and superficial relevance), and IRC will locate all contextually relevant interpretations. The disjunction of the set of contextually relevant interpretations from the set of relevant interpretations will leave the superficially relevant interpretations which can be recommended to the user as "See Also" links.

The maximum contextual distance (search depth) and a relevance threshold can be set by the user.

6.6.1 Determining contextual relevance

In Section 6.5 we suggested that contextual distance between two interpretations influences the degree of contextual relevance. At a particular point (for example, if more than ten nodes separate them) even two interpretations on the same context path will cease to be contextually relevant, and will become superficially relevant or non-relevant. Through contextual distance there seems to be scope for defining a weaker notion of a context switch which allows the transition from one interpretation to another which is not strictly contextually relevant, given the current definition, but which is superficially relevant and meets some additional criteria.

Consider the situation where AID recommends to the user, via a "See Also" link, a superficially relevant interpretation. The user decides to access the interpretation. As there is no context path from the node the user is current visiting, or the number of intervening nodes is too great, the user cannot be guided to the superficially relevant interpretation, and must, instead be taken there directly. Strictly speaking, a context switch has occurred, and the current context session should be replaced by a new, initialised session rooted by the newly accessed interpretation. However, this is counter-intuitive, because the superficially relevant interpretation was located because it is consistent with the user's interests and the user has also decided to visit it. The superficially relevant interpretation should be an extension of the current context session. We also do not distinguish between a superficially relevant interpretation which is accessible from the current interpretation only through the context bottom (and which is consequently the greatest possible contextual distance from the current interpretation), and an interpretation which is a sibling of the current interpretation (see figure 6.1).

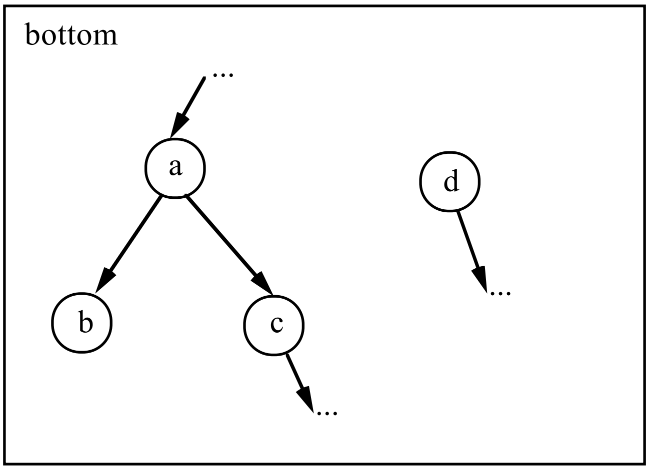

Figure 6.1: Siblings

Again, intuition suggests that the interpretation at node b is considerably more contextually relevant to node c than to node d, which is accessible only through the context bottom. Intuition also tells us that we should be able to guide the user to node b from node c, because it involves revisiting a node the user has already visited in this context session (the two interpretations share a common super-context, which is the context in which node a was interpreted). How should we treat nodes which are more distantly related than siblings, but which still share a common ancestor (the point at which two context paths intersect)? These questions deserve answers, as they are likely to considerably enrich HyperContext and also allow HyperContext to add context to more formal network structures such as semantic networks and conceptual graphs. However, we do not attempt to answer these questions in this thesis.

We currently take the rather simplistic view that if a user wishes to access a superficially relevant interpretation, the access will be direct and it will mark the beginning of a new context session. That it will be the start of a new context session is not as great a disadvantage as it might at first seem, because the superficially relevant interpretation is similar to the user query extracted from the model of the user's short term interests. Also, if the user considers that the document is relevant, and if it would benefit future users, she can always create a link to it from the interpretation she was visiting previously.

Although there is considerable scope to extend the notions of contextual distance, contextual relevance and context switching, we currently consider an interpretation to be contextually relevant if it is relevant to a user query and it exists within a certain number of link traversals on the same context path from the user's location. An interpretation is superficially relevant if it is relevant to the query, but it either does not exist on the same context path as the interpretation the user is visiting, or else too many link traversals are required to reach it. Accessing an interpretation through the context bottom is considered to require the greatest possible number of link traversals. Finally, an interpretation is non-relevant if it is not relevant to the query, and it is contextually distant from the interpretation the user is currently visiting. A context switch occurs if the user hyperleaps to an interpretation which is contextually distant.

6.6.2 Top-down versus bottom-up context resolution

The description we have given so far to determine if two interpretations exist on the same context path implies a top-down approach where all links from an interpretation are followed to see if there is a context path from that interpretation to the other. This also implies that the relationship between the interpretations is uni-directional, and that one interpretation must be a descendant of the other. A bottom-up approach would involve first finding all relevant interpretations (through Traditional Information Retrieval, which does not distinguish between contextual and superficial relevance), and then expanding the contexts of each relevant interpretation to identify those which intersect with the interpretation the user is currently visiting. This is possible because each interpretation knows which contexts supports it. The context of the relevant interpretation can be expanded to all interpretations of node which contain label linked to the document supporting the relevant interpretation. Basically, this is tantamount to searching for all node-label-document links, where node-label is the context of the relevant interpretation and document is the name of the document upon which the relevant interpretation is based. This expansion continues until an interpretation is reached which is an interpretation of a document in the context bottom, or the contextual distance threshold for contextual relevance is exceeded, or the interpretation the user is currently visiting is encountered in the expanded context path.

The interpretations that are contextually close to an interpretation form the interpretation's *context sphere*, and the number of interpretations contained within a context sphere are finite. Given that a top-down and a bottom-up method for determining contextual relevance will yield identical results, we wish to identify the more computationally economical method for a particular situation. Although Traditional Information Retrieval will give the precise number of relevant interpretations which exceed the relevance threshold, it does not necessarily mean that the same number of context paths will need to be expanded, because it is possible for different context paths to contain the same interpretation of a document - many interpretations of the same document may contain a link on the same label to the same destination. The question of whether a top-down or bottom-up method of context expansion is more economical for a given situation is a open problem. However, we recommend a parallel distributed solution for which the client is the co-ordinator.

6.7 Context and the short-term user model

Chapter 5.8 describes a three-stage process to update the model of the user's short-term interests based on the user's path of traversal, and the derivation of a query from the user model. The first stage (Chapter 5.8.1) involves comparing the interpretation the user is currently visiting to other interpretations of the same document, to synthesise the user's interests as a *salient interpretation*. In the second stage (Chapter 5.8.2), the salient interpretation is incorporated into the user model. A query is abstracted from the model of the user's short-term interests in the third and final stage of the process (Chapter 5.8.3). In this section we describe the processes involved in updating the user model in terms of context.

6.7.1 Deriving a salient interpretation

In Chapter 5.8.1, we describe a method to abstract information about a user's short-term interests based on a single interpretation. Ideally, the user indicates those parts of the document which are relevant to her information seeking task, and the extent to which they are relevant. In HyperContext we automatically determine this information.

An interpretation is represented as a vector of weighted terms which jointly describe a document in context. For a user who accesses an interpretation, but who does not end the information seeking task at that interpretation, the interpretation is only partially relevant to her information need. We want to determine which terms used to describe the interpretation are actually relevant and to what degree.

When a user creates an interpretation of a document in context, we can consider that the interpretation is a (partial) representation of the user's cognitive context. When a user accesses an interpretation, we attempt to represent the user's cognitive context in the short-term user model. We then assume that the user is interested in information (in context) which is admitted by the user's cognitive context. If the user terminates the context session at an interpretation and indicates that the interpretation satisfies her information need, we would hope that the short-term user model would have indicated that interpretation to be relevant at some stage *prior* to the user accessing the interpretation. This would indicate that the user's cognitive context was adequately represented in the model of the user's short-term interests *and* that the user model *admitted* the interest represented in the interpretation of the satisfying document. A failure at either point would result in HyperContext misleading, or at least misrepresenting, the user. The results of experimental studies are presented in Chapter 9, but for the remainder of this section, we will rationalise the approach taken to estimating and representing the user's interest with respect to information in context.

The first step in estimating a user's interest is based on an analysis of the interpretation of a document she has just accessed. Assume that the document is at least partially, but not totally, relevant to her information seeking task, in the sense that it contains information in partial fulfilment of her task or she feels that the document can lead to relevant information. The user can counteract this assumption by giving direct feedback that the document is relevant or irrelevant. If the document contains precisely the information, or the last piece of information, she is looking for, then the context session will terminate, ending the requirement to estimate her short-term interest. If the context session does not end, then the user's interest in the document can be estimated in at least three different ways. The first is to assume that the accessed interpretation sufficiently represents the user's interest in the document. The second approach is to assume that the user is interested in the *pertinent features*, or distinguishing characteristics, of the interpretation. These can be identified by comparing the interpretation with the average interpretation of the document. The third and final approach that we consider is the converse of the assumption underlying the second approach. Here we assume that if the accessed interpretation is not the end of the context session, then its pertinent features can distract attention from the actual user interest. There are many other approaches which can be used to gauge the user's particular interests in the current document. One such approach, which we have not investigated, is to compare the interpretations of the selected interpretation's siblings (if any). In this case, we again assume that the user is interested in the interpretation's pertinent features, but we determine them by comparing the selected interpretation with its siblings. We call the resultant synthesised interpretation the *salient interpretation*.

Irrespective of the selected approach to estimating the user's interest in the current interpretation, we will ultimately incorporate the user's estimated interest into the short-term user model. This step is described in greater detail in Section 6.7.2, but for now we remind the reader that the user model will be used to identify the information the user seeks. Although we know what information the user has seen, we must make educated guesses about the relevance of the information the user has not yet seen. This entails that a method of deriving information which returns a document that the user has already seen in the current context session is less suitable than one which consistently returns a document that the user has not already seen. Although returning a random document would meet this criterion, we reject it as the process must be predictable, reproducible in similar circumstances, and we require the document we return to the user to be relevant to the user query.

When a user extends a context session beyond a particular interpretation (by, for example, following a link to another document) we can assume that the interpretation contains at best only partial information about a user's potential interests. Our task is to determine what information about the interpretation is relevant and what is not relevant. However, although the user has accessed an interpretation (which itself may be only a subset of the "complete" description of the document) the user has seen the *entire* document. We have conducted a series of experiments to determine an effective approach to deriving a salient interpretation in order to identify a relevant document which the user has not already seen during the context session. The process and results are described in detail in Chapter 9, but the method which consistently returned relevant documents the user had not already seen (57.3% of the time) was interpretation-, which assumes that the accessed interpretation's pertinent features can actually distract attention from what the user is interested in. The next best method, plain, is based on the current interpretation without reference to other interpretations of the same document, and recommended an unseen document 47.1% of the time. These two methods far out-performed interpretation+, which assumed that the user is *more* interested in the pertinent features of the current interpretation (20%), and control, the control method (20.4%). The control method did not utilise different interpretations of documents, but instead used the same representation of a document in each of its contexts (the equivalent of a single index per document). It is important to note that although both the interpretation- and plain methods returned a high percentage of unseen documents, they each tended to recommend different documents most of the time. However, for this experiment, we were more interested in ensuring that the recommended documents were relevant and unseen. A second series of experiments were conducted to obtain user-provided relevance judgements on different documents recommended by interpretation- and control (Chapter 9.7).

It is, perhaps, counter-intuitive that the user is assumed to be less interested in the pertinent features of the accessed interpretation than in the unaccessed interpretations of the same document. After all, a user's path of traversal through interpretations of documents is intended to be a reflection of her interests. Although the method based solely on the accessed interpretations (plain) also performed reasonably well, the following example may provide an insight into the appropriateness of interpretation-.

Consider the following conversation between two people. As the conversation progresses, it dawns on one of them that they have been talking at cross-purposes:

Mark: "Do you like squash?"
Diane: "I enjoy it from time to time, but it's not a favourite of mine."
Mark: "Oh? Why not?"
Diane: "I prefer something that's not quite so hard"
Mark: "Squash isn't hard - at least, not if it's done properly."
Diane: "I don't think so. It seems that the better you are at it, the more gruelling it becomes."

At this point in the conversation, Mark is surprised by Diane's description of squash being gruelling, and realises that she is talking about "Squash Rackets"[3] whereas he has been talking about "squash", the marrow-like vegetable. Although the reasons for the misunderstanding, and for Mark to realise that he and Diane are not discussing the same thing, can perhaps be easily explained, it is perhaps more difficult to explain how Mark realises that Diane believes the topic of conversation is "Squash Rackets". From the literature review on context (Section 6.2), each person uses or constructs a cognitive context within which to "understand" the conversation. Each person's cognitive context

[3] Squash Rackets is the official name of the sport more commonly known as squash.

usually includes a representation of what the person believes about the other person's cognitive context. Communication failure can occasionally occur when one of the participants misunderstands the other. One of the ways in which a misunderstanding may occur may be due to the listener using a set of assumptions to interpret an utterance which are inconsistent with the set of assumptions that the speaker used to generate the sentence. Consider that Mark believes that Diane believes that the conversation is about "squash: the vegetable". Conversely, when Mark utters the sentence "Do you like squash?", Diane believes that Mark believes the topic of conversation to be "Squash Rackets". Mark is surprised by Diane's last utterance because "gruelling" is not a concept which his cognitive context associates with "squash: the vegetable", but, as predicted by Grice's Laws of Conversational Implicature [42], he is co-operative and so he attempts to work out why "gruelling" was used. Even if Mark does not know about "Squash Rackets" he can tell that something is not quite correct about Diane's response, and may elicit a direct response by asking what she means by it. However, Mark may be aware of "Squash Rackets" as a concept which admits the description "a gruelling sport". However, there is nothing in the context of "squash: the vegetable" which enables Mark to infer that the conversation is about "Squash Rackets". Mark needs to transcend his cognitive context, to recognise that the pertinent features of the current interpretation of "squash" are those which prevent a more appropriate interpretation from being realised, and to relax those assumptions to allow a transition to another context in which all of Diane's utterances can be supported. As the plain and the interpretation+ methods are introspective and self-promoting respectively, from the point of view of reasoning in context, it is reasonable to consider it harder to transcend contexts with them because those methods try to justify "misconstruals" and "misunderstandings" in terms of the interpretation's own description, rather than of in terms of the descriptions of other interpretations of the same document.

In HyperContext then, the salient interpretation is obtained with reference to other interpretations of the same document. We relax the assumptions of the current interpretation of the document in context by introducing features from other interpretations of the same document in which the user could be interested, based on the frequency with which terms generally occur in the other interpretations.

6.7.2 Incorporating the salient interpretation into the user model

The salient interpretation gives us an impression of what the user's interests might be based on the analysis of interpretations of a single document. The salient interpretation must be incorporated into the user model to yield a more accurate representation of the user's interests based on the overall context session.

In Chapter 5.8.2, we describe how salient interpretations are weighted according to the confidence we have in them prior to being combined into a user model. The longer a context session lasts, the greater the confidence we have in the salient interpretation of the currently accessed interpretation, and in terms which occur frequently in the salient interpretations over the context session. This reflects our belief that the longer a user follows a context path, the more likely it is that we can accurately estimate the user's short-term interests, and that unless features occur regularly during a context session, those which occur later in the context session should have more of an influence on the user model than those which occur earlier. In Chapter 5.8.2, we weight the salient interpretations on a confidence scale of 0.125, 0.25, 0.3, 0.4, 0.6, 0.7, and 0.8, for the first, second, third, to seventh and subsequent salient interpretations respectively. Consequently, terms which occur in earlier salient interpretations only will have less influence than those which occur only in the later ones, but terms which occur regularly throughout a context session will have a greater influence than those which occur only in the latter stages.

Our experiment measured the frequency of unseen document recommendations with and without first weighting the salient interpretations. The results referred to in Section 6.7.1 include the salient interpretations weighted according to this scale. In parallel, we also derived performance data for unweighted salient interpretations. In the control and interpretation+ methods, a slight increase in the number of recommendations of unseen documents was observed (from 20.4% to 22% and from 20% to 23.7% respectively), whereas a decrease was observed in the plain and interpretation- methods (from 47.1% to 42.2% and from 57.3% to 51.1% respectively). Other pertinent data collected during the experiment suggest that the combination of methods which most regularly recommends previously unseen documents is the weighted interpretation- method. This data is presented in Chapter 9.

6.8 Summary

People who participate in a conversation may construct cognitive contexts in which they understand the conversation and represent their beliefs about the other participants. Context may be represented in computer-based domains to improve understanding between interacting sub-systems, including users.
HyperContext uses context to partition hyperspace, describe information in context, identify a user's interests, and to guide users to relevant information. There is a high degree of dependency between a context, a link, and the interpretation of the document at the destination of the link. A link is represented by the triple node-label-document,

and the context node-label is extracted from the link during link traversal to interpret the document.

We distinguish between *contextual* and *superficial* relevance, as a function of relevance of an interpretation to a user query and the *contextual distance* between the relevant interpretation and the interpretation the user is currently visiting. If these two interpretations are located within the same *context sphere*, implying that the contextual distance between them is small, they are considered to be contextually relevant. If they are contextually distant they are considered to be superficially relevant. Traditional Information Retrieval will locate all interpretations that are similar to a query, regardless of contextual distance, whereas Information Retrieval-in-Context locates those interpretations which are contextually nearby. Adaptive Information Discovery recommends superficially relevant interpretations via "See Also" links. The user can be guided to contextually relevant interpretations.

Adaptive Information Discovery generates a query from a user model of the user's short-term interests. We justify, in terms of context, our approach to deriving the salient interpretations of documents on which the user model is based.