Temporal Interpretation, Discourse Relations and Common Sense Entailment*

Alex Lascarides[†]
Human Communication Research Centre
University of Edinburgh

Nicholas Asher
IMS, University of Stuttgart, and
University of Texas, Austin

Abstract

This paper presents a formal account of how to determine the discourse relations between sentences in a text, and the relations between the events they describe. The distinct natural interpretations of texts with similar syntax are explained in terms of defeasible rules. These characterise the effects of causal knowledge and knowledge of language use on interpretation. Patterns of defeasible entailment that are supported by the logic in which the theory is expressed are shown to underly temporal interpretation.

1 The Problem of Temporal Relations

An essential part of text interpretation involves calculating the relations between the events described. But sentential syntax and compositional semantics alone don't provide the basis for doing this. The sentences in (1) and (2) have the same syntax, and so using compositional semantics one would predict that the events stand in similar temporal relations.

(1) Max stood up. John greeted him.

(2) Max fell. John pushed him.

But in (1) the order in which the events are described *matches* their temporal order, whereas in (2) descriptive order *mismatches* temporal order. At least, these are the natural interpretations of (1) and (2) unless the reader has information to the contrary. Similarly, the syntax of the sentences in texts (3) and (4) are similar, but the natural interpretations are different.

(3) Max opened the door. The room was pitch dark.

(4) Max switched off the light. The room was pitch dark.

*We would like to thank Mimo Caenepeel, Hans Kamp, Marc Moens, Michael Morreau, Jon Oberlander and an anonymous reviewer for their helpful comments.

†The support of the Science and Engineering Research Council through project number GR/G22077 is gratefully acknowledged. HCRC is supported by the Economic and Social Research Council.

The event and state in (3) temporally *overlap*, whereas in (4) they do not. An adequate account of temporal interpretation must capture these distinctions.

2 The Goals

Previous treatments of tense in Discourse Representation Theory (DRT) have investigated temporal relations in *narrative* texts like (1), (3) and (4) (Kamp & Rohrer 1983, Partee 1984, Hinrichs 1986). Dowty (1986) has explored the semantics of narrative text in an interval-based framework. These theories characterise tense as a tripartite relation between event time, reference time and speech time. The reference time is anaphoric (cf. Reichenbach 1947): the forward movement of time in (1) is encoded in the logical form of the clauses through the forward movement of their reference times. In the case of DRT, statives don't invoke forward movement of reference times, and thus a distinction between (1) and (3) is achieved.

None of these theories attempt to account for the apparent backward movement of time in (2), and in texts like (5).

(5) The council built the bridge. The architect drew up the plans.

Moreover, they are unable to explain why the natural interpretations of (3) and (4) are different. We aim to extend these theories to solve the above problems. We will show that temporal relations must be calculated on the basis of semantic content, knowledge of causation and knowledge of language use, as well as sentential syntax and compositional semantics. Consequently, temporal interpretation will *not* be determined by relations between reference times, where those relations are encoded in a logical form built from syntax alone. This is in sharp contrast to the traditional DRT interpretation of tense.

We'll argue that the basis for distinguishing (1) and (2) is a piece of defeasible causal knowledge relating falling and pushing, which is lacking for standing up and greeting: the knowledge is defeasible in the familiar sense that it is subject to exception. More generally, we'll suggest that defeasible reasoning underlies the interpretation of text: A defeasible logic is one where, informally, Γ defeasibly implies ϕ just in case from knowing only Γ one would infer ϕ. Thus conclusions following from a set of premises need not follow from a superset. We'll show that this type of logic provides a suitable system of inference for modelling the interactions between the Gricean pragmatic maxims and the world knowledge used to calculate temporal structure during interpretation.

Fixing NL interpretation in a defeasible reasoning system has been proposed before. In the realm of temporal interpretation, it has been defended by Hobbs (1979, 1985) and Dahlgren (1988). In certain respects, the approach developed in the current paper refines that outlined by Hobbs and Dahlgren. They encode world knowledge (WK) and linguistic knowledge (LK) in a declarative framework, and use this to determine the preferred interpretations of text. Our approach is similar in this respect. The main difference, in fact, lies in the utilization of a constrained representation of defeasible knowledge, in which the underlying relation of logical consequence yields the interactions required.

In Hobbs' and Dahlgren's theories as they stand, it is not clear that the requisite notion of logical consequence could be defined, since the laws that need to interact in certain specific ways are not related in any obvious way. In particular, conflict can arise among the knowledge sources that they recruit during interpretation, and on occasion this conflict must be resolved. But since the resolution of conflict lacks logical justification, the interaction among the various knowledge sources appears arbitrary. We maintain that the only way of putting the reader's knowledge to work is to place it in the context of a logical language, for which a notion of logical consequence can be defined. Unlike Hobbs and Dahlgren, we will place the reader's knowledge in a logic where its implications can be precisely calculated.

(Lascarides 1992) and (Lascarides & Oberlander 1992) calculate temporal structure by using WK and LK represented in terms of defeasible laws. But they dealt only with the temporal structure derived from two *consecutive* sentences in a text. The theory presented here can be viewed as an extension of that in (Lascarides & Oberlander 1992) in several ways. First, we supply a logic that supports the inferences they discuss. Second, we suggest an analysis of the pluperfect that explains why (1) and (6) are different.

(6) Max stood up. John had greeted him.

Finally, we calculate the relation between sentences separated by intervening material, such as the relation between (7a) and (7e) in the text below:

(7) a. Guy experienced a lovely evening last night.
b. He had a fantastic meal.
c. He ate salmon.
d. He devoured lots of cheese.
e. He won a dancing competition.

In order to extend coverage to texts like (7) we assume, in line with (Hobbs 1985, Grosz & Sidner 1986, Thompson & Mann 1987, Scha & Polanyi 1988), that constraints must be imposed on which sentences in a text can be related to form text segments, and that these constraints are to be characterised in terms of *hierarchical discourse structure*. We therefore not only calculate the temporal structure of the events described in a text. We also investigate how WK and LK affect the interactions between discourse structure and temporal structure within very simple discourses.

The basic model of discourse structure we explore is one where units of a discourse are linked by discourse or rhetorical relations modelled after those proposed by Hobbs (1985). These discourse relations determine the hierarchical structure of the discourse, and hence the structural constraints on which sentences can attach together to form text segments. Because we're concerned only with temporal aspects of interpretation, we consider here only certain discourse relations that are central to temporal import. These are listed below, where the clause α appears in the text before β:[1]

[1] Extensive classifications of discourse relations are offered in (Hobbs 1985, Thompson & Mann 1987, Scha & Polanyi 1988, Asher in press).

- *Explanation*(α, β): the event described in β explains why α's event happened (perhaps by causing it); e.g., text (2).
- *Elaboration*(α, β): β's event is part of α's (perhaps by being in the preparatory phase); e.g., text (5).[2]
- $Narration(\alpha, \beta)$: The event described in β is a consequence of (but not strictly speaking caused by) the event described in α; e.g., text (1).
- $Background(\alpha, \beta)$: The state described in β is the 'backdrop' or circumstances under which the event in α occurred (no causal connections but the event and state temporally overlap); e.g., text (3).
- $Result(\alpha, \beta)$: The event described in α caused the event or state described in β; e.g., text (4).

Explanation is in a sense the dual to *Result*; they both invoke causation, but the latter matches the textual and temporal orders of the events whereas the former doesn't. Both *Result* and *Narration* encode that textual order matches temporal order, but only the former relation induces a causal link between the events. We will provide an account of how these relations constrain the hierarchical structure of text, and we will supply logical mechanisms for inferring the relations from the reader's knowledge resources.

The theory presented here lies at the intersection between formal semantics and computational linguistics. The relevance to formal semantics lies in our extension of existing DRT treatments of tense to solve the problems described. The extension will involve augmenting the DRT framework with the rhetorical relations mentioned above. We use this DRT-based theory of discourse structure so that direct comparisons with theories of tense in formal semantics can be made. But our account will be simpler than previous DRT treatments, in that we will not consider the role of temporal connectives in structuring discourse. With regard to computational linguistics, we refine the proposal that defeasible reasoning underlies NL interpretation. We do this by examining in detail how a well defined notion of nonmonotonic consequence can be used to calculate precisely the interactions between semantic content, causal knowledge and pragmatic maxims. However, we offer no implementation, and we largely ignore communicative goals and intentions (cf. Grosz and Sidner 1986, Cohen, Morgan and Pollack 1991).

3 The Basic Story

A logic will be suitable for calculating temporal interpretation only if it supports all the patterns of inference we need to calculate temporal and rhetorical relations. These inferences are listed below; we motivate each one by a particular example. For simplicity, most of these examples will involve inferring temporal relations rather than discourse ones; the latter will be dealt with upon full formalisation.

[2] We assume Moens and Steedman's (1988) tripartite structure of events, where an event consists of a preparatory phase, a culmination and a consequent phase.

Some formal notation is useful in order to clarify the logical structure of the inferences. We assume $A \approx\!\!\!\mid B$ represents "B nonmonotonically follows from A" and $A \models B$ represents "B monotonically follows from A". $\phi > \psi$ is a gloss for the default rule "if ϕ, then normally ψ" and $\phi \rightarrow \psi$ is a gloss for "if ϕ, then (indefeasibly) ψ". We should stress that this formal notation is intended to be theory-neutral at this stage, for we simply wish to survey which patterns of inference we require from the logic. Thus for now, one can consider $\phi > \psi$ to be syntactic sugar for the way defaults are represented in default logic (i.e. $\phi, M\psi/\psi$, where M is the consistency operator), in autoepistemic logic (i.e. $L\phi \wedge \neg L \neg \psi \rightarrow \psi$, where L is the modal autoepistemic operator), or in circumscription (i.e. $\phi \wedge \neg abnormal_n(\phi) \rightarrow \psi$, where $abnormal_n$ is an abnormality predicate). The two notions of entailment, i.e. $\mid\!\approx$ and $\models$, can equally be regarded as syntactic sugar for the different notions of entailment supported in these various logics.

3.1 Defeasible Modus Ponens

Defeasible Modus Ponens is the following nonmonotonic inference:[3]

- **Defeasible Modus Ponens**
  $\phi > \psi, \phi \mid\!\approx \psi$
  E.g., Birds normally fly, Tweety is a bird $\mid\!\approx$ Tweety flies

We will show how it underlies the interpretation of text (1).

Intuitively, it's sometimes the case that the *only* information available to an interpreter regarding temporal structure is textual order. This is the case when no information about the temporal order of events is derivable from WK and LK, linguistic context, or syntactic clues like a change in tense, words like *because*, or adverbials. In such cases we claim the descriptive order of events typically matches temporal order. As explained in (Lascarides 1992) and (Lascarides and Oberlander 1992), this claim is in line with Grice's (1975) Maxim of Manner, where it is suggested that text should be orderly. One way of interpreting this is that events should be described in the right order. Furthermore, Dowty (1986) claims that in some genres at least, the author typically describes events in the order of their discovery or perception, resulting in descriptive order matching temporal order. Our claim is that this LK plays a key role in interpreting (1). This is because in (1) it is indeed the case that the only information available for calculating temporal structure is textual order.

We represented this temporal information conveyed by textual order as a combination of two rules; one defeasible and one indefeasible.

- **Narration**
  If the clause β currently being processed is to be attached by a discourse relation to the clause α that's part of the text processed so far, then normally, $Narration(\alpha, \beta)$ holds.

- **Axiom on Narration**
  If $Narration(\alpha, \beta)$ holds, and α and β describe the eventualities e_1 and e_2 respectively, then e_1 occurs before e_2.

[3] We ignore quantification and variables here, because the logic we ultimately use will be propositional.

Narration exploits the dynamic processing of discourse familiar in DRT, where text structure is built by processing the successive sentences: since β is to be attached to α, α must have been in the text before β, and thus the textual order of α and β is coded in Narration's antecedent. Thus Narration and its Axiom together convey the information that textual order matches temporal order, unless there's information to the contrary. The idea that Gricean-style pragmatic maxims should be represented as defeasible rules has already been suggested in (Joshi, Webber and Weischedel 1984). Here, we've argued that the defeasible law Narration is a manifestation of Grice's Maxim of Manner. It is defeasible because there are exceptions, e.g., text (2).

Suppose that the logical forms of the sentences in (1) are respectively α and β; they describe the events e_1 of Max standing up and e_2 of John greeting him, but they do *not* impose any conditions on the relations between e_1 and e_2. Then Narration and its axiom can contribute to the interpretation of (1), providing the reader's KB contains the following:

(a) α and β (so the reader believes the author is sincere)
(b) β is to be attached to α by a discourse relation (so the reader believes the text is coherent)
(c) All defeasible WK, such as causal laws, and all defeasible LK, such as Narration.
(d) All indefeasible knowledge, like the Axiom on Narration and that causes precede effects.
(e) The laws of logic.

Defeasible Modus Ponens captures the preferred reading of (1) with respect to this KB:

(i) The premises of Narration are verified by the KB (assumption (b)).
(ii) The consequent of Narration is consistent with the KB.
(iii) (i) and (ii) form the premises of Defeasible Modus Ponens; $Narration(\alpha, \beta)$ is inferred.
(iv) Therefore, by the Axiom on Narration, Max standing up precedes John greeting him.

It must be stressed that this line of reasoning doesn't represent a psychologically plausible account of human text processing. Assumption (e) alone prevents us from making such claims. Nonetheless, the above reasoning predicts the correct analysis of (1) with respect to the reader's knowledge about the context, language use and the world.

3.2 The Penguin Principle

The Penguin Principle is the following pattern of inference:

- **Penguin Principle**
  $\phi \rightarrow \psi, \phi > \neg\chi, \psi > \chi, \phi \mathrel{|\!\!\approx} \neg\chi$
  E.g., Penguins are birds, penguins normally don't fly, birds normally fly, Tweety is a penguin $\mathrel{|\!\!\approx}$ Tweety doesn't fly

The antecedent to two defeasible laws are verified, and the conclusions to both cannot hold in a consistent KB. In this sense, the rules conflict. The antecedent to one of the laws entails the antecedent to the other; in this sense, it is more specific. We want a defeasible logic to validate the inference in which the more specific rule is preferred. It will play a central role in resolving conflict among the reader's knowledge resources.

We now show how the Penguin Principle contributes to the interpretation of (2). First, we describe the relevant defeasible rules. By assumption (b) on the KB, an interpreter will infer a discourse connection between the sentences in a two-sentence text like (2). From the way we understand the discourse relations to which we limit ourselves in this paper, a discourse relation between two clauses entails one of the following relations between the *eventualities* (i.e. events and states) described: either one is a consequence of the other, one is part of (i.e. in the preparatory phase of) the other, they stand in causal relations, or they temporally overlap. For brevity, we state that if one of the above relations holds between two eventualities e_1 and e_2, then e_1 and e_2 are *e-connected*. Since the reader infers that the sentences in (1) and (2) are discourse related, he infers that the events they describe are e-connected. In the case of (2), there is WK gained from perception and experience, that relates falling and pushing in a particular way:

- **Push Causal Law**
  If e_1 where x falls and e_2 where y pushes x are e-connected, then normally, e_2 causes e_1.

There is also indefeasible knowledge that causes precede effects:

- **Causes Precede Effects**
  If e_1 causes e_2, then (indefeasibly) e_2 does not precede e_1.

The Push Causal Law is a *defeasible* law. Unlike Dahlgren's (1988) probabilistic laws, it should *not* be read as pushings usually cause fallings. Indeed, we would claim that such a law would be far-fetched. For there will be plenty of pushings that don't cause fallings; and there may well be plenty of fallings that cause pushings. Rather, the Push Causal Law states that *if* a pushing and falling are e-connected, *then* although both causal directions may be permissable, one normally prefers one to the other.[4]

The Causal Law may seem very 'specific'. It could potentially be generalised, perhaps by re-stating e_1 as x moving and e_2 as y applying a force to x, and then having appropriate connections between pushings and applying force, and fallings and movement to interpret (2). We're concerned, however, with making the case for exploiting WK and LK in a logical framework to drive temporal interpretation, and this case can be made while glossing over generalisations like these. So we do so here.

There is no similar law for standing up and greeting; if one knows that a standing up and a greeting are e-connected but one knows nothing more, then it's not possible to conclude, even tentatively, exactly how the events are related. Our claim is that the Push Causal Law forms the basis for the distinction between (1) and (2).

[4] We are not concerned here with the metaphysics or meaning of causality. Suffice to say for now that our representation of causal relations is compatible with many of the current theories on causality. We also leave open the question of how these causal laws are acquired; suffice to say that they do not represent implausible knowledge.

In interpreting (2), there is conflict among the applicable knowledge resources, specifically between Narration and the Causal Law; given Narration's Axiom and Causes Precede Effects, the consequents of both laws can't hold in a consistent KB. Given that the natural interpretation of (2) in the 'null' context is one where the pushing caused the falling, this conflict must be resolved in that context in favour of the Causal Law.

In nonmonotonic logics, conflict between defeasible rules is resolvable if one is more specific than the other, cf. the Penguin Principle. Given that clauses related by a discourse relation describe events that are e-connected, we can restate the Push Causal Law as follows, making its antecedent entail that of Narration:

- **Push Causal Law**
  If β is to be attached to α, and α describes an event e_1 of x falling and β describes an event e_2 of y pushing x, then normally, e_2 causes e_1.

The new version of the Push Causal Law expresses a mixture of defeasible WK and LK, for it asserts that given the sentences are discourse-related somehow there is a e-connection between the events, and given the kinds of events they are, the second event described caused the first, if things are normal.

This new statement of the Causal Law ensures that together with Narration and the above contents of the reader's KB, the premises of a Penguin Principle are formed. α and β are now respectively the logical forms of the sentences in (2) rather than (1); the antecedents to the Causal Law and Narration are verified; these rules conflict and the more specific rule—the Causal Law—wins, resulting in an inference that the pushing caused the falling. This deals with the temporal structure of (2). We'll subsequently show that inferring the discourse structure of (2) also involves a Penguin Principle; in fact the inferences about temporal structure and discourse structure will be interleaved during interpretation.

To ensure that the logic resolves the conflict in the right way when interpreting (2) with respect to the above context, we have essentially had to tailor the representation of causal knowledge to the particular task at hand, namely linguistic interpretation. For we have had to explicitly represent how the information is presented in the text in the 'causal' rule, in order to achieve the right logical interactions with other knowledge resources. Causal laws represented as "an event e_1 of type a normally causes an event e_2 of type b" would not bear the right logical relation to the available LK concerning textual order.

3.3 The Nixon Diamond

Conflict among defeasible rules is not resolvable. when the antecedents of the rules aren't related. We refer to this (lack of) inference as the Nixon Diamond:

- **Nixon Diamond**
  $\phi > \chi, \psi > \neg\chi, \phi, \psi \not\approx \chi$ (or $\neg\chi$)
  Quakers are normally pacifists, Republicans normally are non-pacifists, Nixon is a Quaker and a republication $\not\approx$ Nixon is a pacifist/Nixon is a non-pacifist.

We argue that the Nixon Diamond provides the key to textual incoherence, as in (8).

(8) ?Max won the race. He was home with the cup.

The appropriate KB when interpreting (8) verifies the antecedent of Narration. In addition, we claim that the following laws hold:

- **Win Law**
  If e_1 is Max winning, and e_2 is Max being at home, then normally, these eventualities don't overlap.
- **States Overlap**
  If the clause β is to be attached to the clause α, and β describes a state s, then normally s overlaps the eventuality e described by α.

The Win Law captures the WK that if Max wins the race and if Max is at home, then these events don't temporally overlap, if things are normal. In other words, it is unusual for the finish line of the race to be at the winner's house. Note that the Win Law doesn't require the antecedent to assert that the event and state are connected, nor in turn does it require the clauses that describe these eventualities to be discourse related. For the intuition this law captures is that the event and state don't normally temporally overlap, *regardless* of whether they are connected or not.

States Overlap, on the other hand, is defeasible LK. This law can be seen as a manifestation of Grice's Maxim of Relevance, as suggested in Lascarides (1992). The argument goes as follows. An NL text describes a portion of the time line, with events and states occurring on it. Let's call this a situation. To construct a full picture of the situation, the interpreter must infer the relative occurrences of the states and the culminations of events, including where states start and stop. Since culminations are punctual (holding at points of time), their relative order is inferred from rules like Narration, which models describing things in the order of perception as explained earlier. But states are extended, and the order of perception of states does not fully determine where the states start relative to the other eventualities. So order of perception is insufficient for determining where a state starts.

There are several linguistic mechanisms which can be used to indicate where a state starts. The author could explicitly refer in the text to what caused the state, and so from the law that causes precede effects, the interpreter will know the relative place where the state starts, cf. text (4). Alternatively, the author can use temporal adverbials to say where a state starts. But (3) shows that these two mechanisms are not enough: how do we determine where states start in text that do *not* feature adverbials or causes? States Overlap is then a vital mechanism for determining where a state starts relative to other eventualities described in text. It basically says that if there is no "explicit" indication of where the state starts—via the mention of causes or the use of temporal adverbials—then the start of the state is assumed to be *before* the situation that the text is concerned with occurs, resulting in overlap given the relation between textual order and order of perception. States Overlap asserts that (unless there's indication to the contrary) the point where a state starts is before the situation that the text describes, rather than part of it; and in that it's not part of the situation, it is irrelevant. Thus States Overlap models the following in accordance with Grice's Maxim of Relevance: the start of a state is relevant only if it is marked as such.

Assume that the reader's KB is as usual, save that α and β are now respectively the logical forms of the sentences in (8). Then the antecedents to States Overlap, the Win Law and Narration are all verified. By the general principle that the more specific default law overrides the less specific one when they conflict, the rule of Narration is deemed irrelevant here; States Overlap is more specific. States Overlap and the Win Law also conflict, but their antecedents aren't logically related. Hence as in the Nixon Diamond, we *fail* to infer any conclusion about the temporal relation between the eventualities described, leading to incoherence.

3.4 Closure on the Right

Closure on the Right is a *monotonic* inference of the following form:

- **Closure on the Right**
  $\phi > \psi, \psi \rightarrow \chi \models \phi > \chi$
  E.g., Lions normally walk, walkers (indefeasibly) have legs $\models$ lions normally have legs.

The theory of discourse attachment we propose requires this inference for several technical reasons, as will be clear in the appendix. It is also used, in conjunction with the Nixon Diamond, to account for discourse popping. Consider again text (7).

(7) a. Guy experienced a lovely evening last night.
b. He had a fantastic meal.
c. He ate salmon.
d. He devoured lots of cheese.
e. He won a dancing competition.

Intuitively, (7c,d) elaborate (7b), but (7e) does not. Thus (7e) should not attach to (7d) to form a text segment. Rather, (7e) should be related by (7b) by *Narration*. But (7d) and (7e) are consecutive sentences, and by our assumptions on the KB they will verify the antecedent to Narration. So something must block Narration's consequent. This is achieved as follows: In line with Scha and Polanyi (1988), we assume that if a text segment is narrative, it cannot also be an elaboration. Thus Rule 1 is part of the reader's KB:

- **Rule 1**
  If $Narration(\alpha, \beta)$ holds, then (indefeasibly) $\neg Elaboration(\alpha, \beta)$ holds.

Rule 1 together with Narration and Closure on the Right produce Rule 2:

- **Rule 2**
  If β is to be attached to α with a discourse relation, then normally, $\neg Elaboration(\alpha, \beta)$ holds.

The following defeasible law captures the intuition that if a clause γ is to attach with *Narration* to a clause β that elaborates α, then γ must also elaborate α:

- **Rule 3**
  If *Elaboration*(α, β) and $\neg$*Elaboration*(α, γ) hold, then normally $\neg$*Narration*(β, γ) holds.

Suppose that the logical forms of (7b), (7d) and (7e) are respectively α, β and γ. Suppose furthermore that in accordance with intuitions, the reader has inferred *Elaboration*(α, β). The task now is to attach γ to the preceding discourse. By Rule 2 with α and γ substituted in the schema, we infer $\neg$*Elaboration*(α, γ). Thus from Rule 3 one would infer $\neg$*Narration*(β, γ). But substituting β and γ in the schema Narration produces the conflicting conclusion *Narration*(β, γ). The antecedents of rules 2 and 3 are not logically related to that of Narration, resulting in irresolvable conflict (cf. the Nixon Diamond). Thus no discourse relation between (7d) and (7e) is inferred, indicating that (7d,e) forms an incoherent text segment. Therefore, (7e) must attach to one of the remaining clauses in the preceding text. A full analysis of this example will be given upon formalisation. For now, suffice to say that Closure on the Right was required to block the attachment of (7e) to (7d).

3.5 Dudley Doorite

Dudley Doorite is also a monotonic inference:

- **Dudley Doorite**
  $\phi > \chi, \psi > \chi \models (\phi \vee \psi) > \chi$
  E.g., Quakers are normally pacifists, Republicans are normally pacifists $\models$ Quakers and Republicans are normally pacifists

We use it to explain the ambiguity of text (9).

(9) The bimetallic strip changed shape. The temperature fell.

The appropriate KB contains the following laws:

- **Change=Bending or Straightening**
  A bimetallic strip changes shape if and only if it bends or straightens.

- **Bend Causal Law**
  If β is to be attached to α, and α describes the event e_1 of a bimetallic strip bending, and β describes an event e_1 of the temperature falling, then normally, e_1 causes e_2.

- **Straighten Causal Law**
  If β is to be attached to α, and α describes the event e_1 of a bimetallic strip straightening, and β describes an event e_1 of the temperature falling, then normally, e_2 causes e_1.

Neither the above causal laws' antecedents are satisfied by the KB in the analysis of (9). However, Closure on the Right and Dudley Doorite can be used to infer the event structure for (9): Closure on the Right yields Law 1 and Law 2 respectively from the Bend Causal Law and the Straighten Causal Law:

- **Law 1**
  If β is to be attached to α, and α describes the event e_1 of a bimetallic strip bending, and β describes an event e_1 of the temperature falling, then normally, e_1 causes e_2 or e_2 causes e_1.

- **Law 2**
  If β is to be attached to α, and α describes the event e_1 of a bimetallic strip straightening, and β describes an event e_1 of the temperature falling, then normally, e_2 causes e_1 or e_1 causes e_2.

These then form the premises for Dudley Doorite which, together with Change=Bending or Straightening, produce the Change Causal Law:

- **Change Causal Law**
  If β is to be attached to α, and α describes the event e_1 of a bimetallic strip changing shape, and β describes an event e_1 of the temperature falling, then normally, e_2 causes e_1 or e_1 causes e_2.

The reader's KB verifies the antecedent to Change Causal Law when interpreting text (9). By Defeasible Modus Ponens, the reader infers that the two events are causally connected, but he fails to infer the *direction* of the causal connection, resulting in temporal ambiguity.

4 Choosing a Nonmonotonic Logic

We will use all the above patterns of inference in text interpretation. So a logic is suitable for our purposes only if the logical consequence relation validates these inferences.

There are several different nonmonotonic logics which handle nonmonotonicity in various ways. For now, we simply survey which logics verify which of the above inferences. Reiter's (1980) default logic and autoepistemic logic (Moore 1984) verify Defeasible Modus Ponens and the Nixon Diamond, but none of the others. Konolige's (1988) Hierarchical Autoepistemic logic (HAEL) verifies Defeasible Modus Ponens and the Penguin Principle, but the premises of a Nixon Diamond yield a contradiction rather than no conclusion. Furthermore, it fails to verify the monotonic patterns of inference. Veltman's (1990) Update Semantics verify all the nonmonotonic patterns of inference, but none of the monotonic ones. It would be difficult to extend these logics to support Closure on the Right (cf. Morreau 1992). Since we require Closure on the Right, these logics are not suitable for our purposes here.

Circumscription (McCarthy 1980) represents defaults using abnormality predicates. For example, *birds fly* is represented as (10), which should be read as: for any x if x is a bird and is not abnormal in sense 1, then x flies.[5]

(10) $(\forall x)(bird(x) \wedge \neg ab_1(x) \rightarrow fly(x))$

[5] We consider here the first order language rather than the propositional version in order to clarify the discussion about abnormality predicates.

The language includes many abnormality predicates, to capture the many senses in which individuals can be abnormal. The default *Penguins don't fly*, for example, is represented in (11); it features a different abnormality predicate to (10) to reflect the intuition that being an abnormal penguin is different from being an abnormal bird.

(11) $(\forall x)(penguin(x) \wedge \neg ab_2(x) \rightarrow \neg fly(x))$

Defeasible entailment in circumscription is defined in terms of those admissable models where the abnormality predicates have as small an extension as is consistent with the premises; the so-called minimal models. ψ defeasibly follows from ϕ if for every minimal model in which ϕ holds, ψ holds. This definition of validity captures Defeasible Modus Ponens and the Nixon Diamond.

Prioritised Circumscription (Lifschitz 1984) captures the Penguin Principle by fixing a particular order to the abnormality predicates, preferring models which minimise over ab_2 to those that minimise over ab_1. Although this captures the intuitively compelling pattern of inference, it only does so in virtue of machinery that's extraneous to the semantics of defaults; namely the ordering of predicates in a hierarchy to guide inference. An alternative way of capturing the Penguin Principle, which does not require this extraneous machinery, is to assume the following law, which states that Penguins are abnormal birds (note we must use the abnormality predicate featured in (10) to express this).

(12) $(\forall x)(penguin(x) \rightarrow ab_1(x))$

The definition of validity in circumscription then captures the intuition that more specific defaults override less specific ones. But this is done at the sacrifice of a uniform translation of NL statements about defaults; for sometimes laws like (12) will have to be encoded in the translation of defaults into the object language.

The fact that NL statement about defaults do not translate uniformly into the formal representation is a general problem with Circumscription; it also raises difficulties when attempting to capture Closure on the Right. Consider the premises of this inference: that is $(\forall x)(A(x) \wedge \neg ab_1(x) \rightarrow B(x))$ and $(\forall x)(B(x) \rightarrow C(x))$. From this it follows that $(\forall x)(A(x) \wedge \neg ab_1(x) \rightarrow C(x))$. But given the multitude of abnormality predicates in Circumscription, we have no way of knowing whether ab_1 is the appropriate abnormality predicate for representing the default connection between A and C. The customary translation of default statements into Circumscription is to introduce a new abnormality predicate for each default statement. So suppose the abnormality predicate to be used when expressing the default connection between A and C is ab_n. Then we fail to support Closure on the Right: the above premises do not imply $(\forall x)(A(x) \wedge \neg ab_n(x) \rightarrow C(x))$. A similar tension arises between capturing Dudley Doorite and translations of NL default statements into Circumscription.

Asher and Morreau's (1991) Commonsense Entailment (henceforward CE) attempts to overcome these problems in Circumscription by restricting all representation of normality to the metalanguage. NL defaults are translated into the formalism in a uniform way. All defaults are represented using a nonmonotonic conditional operator $>$. Intuitively, $\phi > \psi$ is read as *If ϕ then normally ψ*, and *birds fly* is represented as $(\forall x)(bird(x) > fly(x))$. *All* default statements are translated in this way, regardless of their specificity. The need for a uniform translation of default statements is paramount when dealing with the complex commonsense knowledge required to reason about discourse.

CE captures all the patterns of inference we have mentioned. Its approach to Closure on the Right is more satisfactory than Circumscription's because multiple abnormality predicates are avoided. Furthermore, CE supports the Penguin Principle without resorting to machinery extraneous to the semantics of defaults, such as prioritising predicates as in Prioritised Circumscription, or forming hierarchies of defaults as in HAEL. We assume that whatever the semantics of default rules, that more specific defaults should override less specific ones when they conflict should be an inherent part of that semantics. We therefore favour the approach to the Penguin Principle taken in CE.

A further reason for using CE concerns the ability to carry out local inferences in a nonmonotonic framework. In general, it's not possible to divide nonmonotonic inference into a chain of deductions based on parts of the context. This is because adding premises to nonmonotonic deductions doesn't always preserve conclusions, making it necessary look at the premises as a whole. In CE, however, it is possible to divide inferences into a chain of subtheories under special circumstances, as shown in the appendix. We will exploit this property upon formalisation, in order to interleave reasoning about temporal structure and discourse structure. We leave open the question of whether chains of local inferences would be possible in Prioritised Circumscription.

We now examine in detail how the ideas presented here can be formally pinned down. To do this, we augment the DRT framework with rhetorical relations, as proposed in Asher (in press); we show how to represent the appropriate LK and WK in CE; and we show how the appropriate temporal structures and discourse structures for NL texts are achieved.

5 Discourse Structure in DRT

The fragment of NL text considered here will be relatively simple, with no quantification, connectives, adverbials or future tense. We concentrate on characterising the pluperfect and simple past tenses alone. We employ an extension of DRT, so that we can compare directly our theory with previous, influential DRT theories of tense.

DRT represents discourse in terms of a discourse representation structure (DRS), which is a pair of sets identifying respectively the discourse entities and conditions on those entities. The individual sentences in this theory will be represented using DRSs. But the representation of text as a whole *extends* the framework of DRT. Text is represented in terms of a *discourse representation pair* (DRP). This consists of a set of DRSs which define the semantic content of the sentences of the text, and a set of conditions on those DRSs, which characterise the discourse relations between the sentences of the text. So in essence, DRPs are the logical forms of the individual sentences plus discourse relations between them.

This kind of representation is a natural extension of the DRT framework: A DRP is a simpler version of an SDRS (or segmented discourse representation structure) as introduced in (Asher in press). It is simpler in the sense that the discourse relations will relate only DRSs (corresponding to sentences) rather than SDRSs (corresponding to segments of text). Discourse relations between sentences rather than segments of text are adequate for the simple fragment considered here, but not adequate in general. For the sake of simplicity, we have chosen not to extend the DRT framework to SDRT here, although such an extension would be straightforward.

The formal definition of DRPs is as follows:

Definition of DRP and DRP Conditions

- If $\alpha_1, \ldots, \alpha_n$ are DRSs, Con a set of DRP conditions on those DRSs, then $\{\{\alpha_1, \ldots, \alpha_n\}, Con\}$ is a DRP;
- If α_1 and α_2 are DRSs and R is a discourse Relation, then $R(\alpha_1, \alpha_2)$ is a DRP condition on α_1 and α_2.
- If α is a DRS, then the DRP $\{\{\alpha\}, \emptyset\}$ is equivalent to α and shall be referred to as such.

For our purposes, discourse relations are two-place predicates whose semantics are defined by the interpretation function in the model. Thus the semantics of DRPs are obtained from the semantics of DRT given in the appendix simply by extending the interpretation function in the appropriate way.

5.1 The Logical Form of Clauses

We now turn to the question of representing clauses as DRSs. The logical form of sentences will introduce discourse entities denoting events and times to capture deictic reference (cf. Partee 1973). But since we are not concerned with NP reference, for simplicity we will assume NPs are represented by constants. We propose that the logical form of the clause (13) is (13′).

(13) Max stood up.

(13′) $[e, t][t \prec now, hold(e, t), standup(max, e)]$

In words, (13′) introduces discourse entities e and t, and asserts that the event e is Max standing up and e holds at a point of time t earlier than *now*.

Contrary to traditional DRT, the semantics of (13′) is given in a modal framework (as CE is modal), with truth defined relative to possible worlds. Moreover, the movement of time through discourse will be determined by defeasible rules for clause attachment rather than by syntactic-based rules for constructing DRSs.

The representation (14′) of the pluperfect clause (14) captures the intuition described in Moens & Steedman (1988): perfects assert that the consequent phase of the event is in force.

(14) Max had taken the bus.

(14′) $[s, t][t \prec now, hold(s, t), s = CS(e), take(max, bus, e)]$

In the above, $CS(e)$ is a term referring to the consequent state of the event e. The semantics of the function CS will be considered primitive for our purposes (but see Lascarides (1988) for a proposed semantics). In words, (14′) states that the consequent state of the event e of Max taking the bus holds at a point of time t that is earlier than the time of speech *now*. Again, unlike the traditional DRT analysis, we don't represent the pluperfect as a tripartite temporal relation, and the movement of time it evokes through discourse is not encoded in the DRS construction rules.

This deals with clauses: texts are represented as DRPs. The problem is to update the DRP: if τ is the DRP representing the first n clauses of the text, and β is the DRS representing the semantic content of the $n+1^{th}$ clause, then τ must be updated with β by adding β to τ's first set and adding discourse relations between β and DRSs in τ to τ's second set. Updating will consist of:

(1) Defining the constraints on the possible sites for attachment of the DRS β (these will be some subset of DRSs in the DRP τ); and

(2) Defining rules for attachment of β at an attachment site.

We now turn to the first task.

5.2 Constraints on Possible Sites for Attachment

We mentioned in section 2 the generally accepted view that only some components of a discourse structure are open in the sense that one may add new material to them as the discourse is processed. To update a DRP τ with the new information given by the clause currently being processed, one has to relate that clause by means of a discourse relation to an *open* DRS in τ, where openness is as defined below. From an intuitive point of view, the definition of openness captures the following: an open clause is either the previous clause (i.e. the last one added to the DRP), or a clause that it elaborates or explains. This is in line with the openness constraints described in Scha and Polanyi (1988) and Asher (in press).

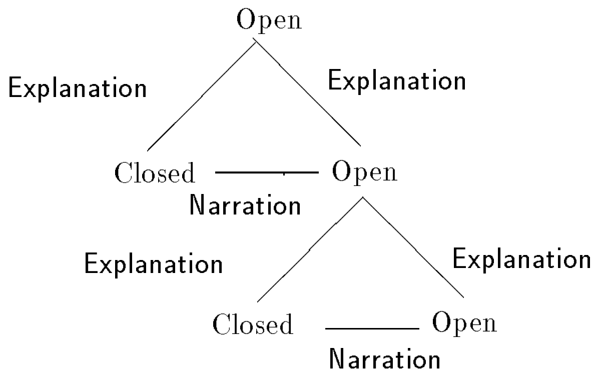
Definitions of Discourse Dominance

Suppose a DRP τ contains the DRSs α and β. Then:

- **Subordination**
  α is *subordinate* to β if:
  (i) $Explanation(\beta, \alpha)$ or $Elaboration(\beta, \alpha)$ holds; or
  (ii) γ is a DRS in τ such that $Explanation(\gamma, \alpha)$ or $Elaboration(\gamma, \alpha)$ holds, and γ is subordinate to β.

- **Openness**
  A DRS α is *open* in the DRP τ if and only if α is the the DRS representing the previous clause in the text, or this DRS is subordinate to α.

One can represent DRPs as graphs, which are built in a depth-first left-to-right manner. The above definition of openness can then be represented pictorially as follows:



There are structural similarities between our notion of openness and Polanyi's (1985) and Webber's (1991); the open clauses are those on the right frontier of the discourse structure.

These constraints on the possible sites for attachment capture two intuitions noted in Scha and Polanyi (1988): first, the only clause to which the current clause can attach in narrative is the previous one; and second, an elaboration or explanation that occurs earlier than the previous clause cannot be resumed.[6]

The awkwardness of text (15) illustrates the first intuition: (15c) should intuitively be related to (15a); but (15a) isn't open since (15a,b) is narrative.

(15) a. Max stood up.
b. John greeted him.
c. ?Max got up slowly.

Text (16) illustrates the second intuition.

(16) a. John arrived late for work.
b. He had taken the bus.
c. He had totalled his car.
d. His boss summoned him to his office.

Intuitively, (16b,c) explains (16a), but (16d) does not. Our definition of openness predicts that (16a-d,e) is awkward (because (16e) attempts to continue the 'closed' explanation (16b,c)), whereas (16a-d,f) is acceptable.

(16) e. ?John had taken his car to the garage.
f. John realised that his chances of promotion looked bleak.

[6] Definite noun phrases that refer to events previously described can occasionally allow one to refer back to segments of text which we claim are 'closed off'. It is debatable, however, whether we re-open closed attachment sites or whether we start a new structure in the discourse. For the sake of simplicity, we ignore these sorts of discourses here, but see (Asher in press).

But we have yet to formally pin down why (16f) is related to (16d) by *Narration*, and not by any of the other discourse relations. We must now define the rules for attaching new information to a particular attachment site. We have argued that these rules must be defeasible, characterising causal laws and pragmatic maxims. We now turn to the problem of representing these defeasible rules within CE.

6 A Brief Introduction to CE

CE supplies a modal semantics for defaults. On top of the truth conditional semantics there is a dynamic partial theory, which accounts for nonmonotonic inferences like those described above. Defeasible Modus Ponens intuitively goes as follows: first one assumes the premises *birds fly* and *Tweety is a bird* and no more than this. Second, one assumes that Tweety is as normal a bird as is consistent with these premises. Finally, from these two assumptions, one concludes that Tweety flies.

Let us first consider the semantics of defeasible statements. As already mentioned, "if ϕ then normally ψ" is represented as $\phi > \psi$. We augment the language to include DRPs, and strict conditionals of the form $\Box(A \rightarrow B)$ in order to represent *indefeasible* laws like the Axiom on Narration. But we also simplify the language in that we use a propositional version.

The semantics of defeasible statements are defined in terms of a function $*$ from worlds and propositions to propositions. Intuitively, $*(w, p)$ is the set of worlds where the proposition p holds together with everything else which, in world w, is normally the case when p holds. So $*$ encodes assumptions about normality. The truth conditions of defeasible statements are defined as follows:

- $M, w \models \phi > \psi$ if and only if $*(w, [\![\phi]\!]) \subseteq [\![\psi]\!]$

In words, the above says that *If ϕ then normally ψ* is true with respect to the model M at the possible world w if the set of worlds that defines what is normally the case when ϕ is true in w all contain the information that ψ is also true. Thus CE is a conditional logic in the Acqvist (1972) tradition; it differs from previous conditional logics in the constraints it imposes on the function $*$. Three constraints are particularly relevant here: Facticity, Specificity and Dudley Doorite, as defined in the appendix. Facticity captures the intuition that however few properties a normal bird has, one of them is that he is a bird. Specificity encodes the constraint that normal birds aren't penguins (because penguins don't have the flying property associated with normal birds). Dudley Doorite encodes the constraint that if Quakers are pacifists and Republicans are pacifists, then Quakers and Republicans are pacifists.

The truth definitions for the full language are given in the appendix, together with the monotonic axioms which make the monotonic part of CE sound and complete. We use a modified version of the logic of CE given in (Asher & Morreau 1991): our axioms reflect 'modal closure' instead of just logical closure, where the background modal theory is S5. We show in the appendix that modal closure is necessary in order to build discourse structures.

Besides the monotonic axioms, CE develops a dynamic semantics of information states, which are sets of possible worlds, that verify defeasible inferences. This makes precise the notion of assuming no more than the premises of an argument (defined in terms of an update function $+$ on information states), and assuming everything is as normal as is consistent with those premises (defined in terms of a normalisation function N). The functions $*$, $+$ and N, which are defined in the appendix, characterise a further notion of (common sense) validity, written $\approx\!\!|$ (also described in the appendix). This supports (at least) Defeasible Modus Ponens, the Nixon Diamond and the Penguin Principle. The monotonic validity $\models$ supports Closure on the Right, and Dudley Doorite (and since $|\!\approx$ is supra-classical, these are also valid under $|\!\approx$). Contrary to HAEL and Prioritised Circumscription, the Penguin Principle is verified without forming hierarchies of defaults.

7 The Dynamic Construction of Discourse Structure

7.1 Narration and Defeasible Modus Ponens

First consider example (1), whose sentences are represented by α and β.

(1) Max stood up. John greeted him.

(α) $[e_1, t_1][t_1 \prec now, hold(e_1, t_1), standup(max, e_1)]$

(β) $[e_2, t_2][t_2 \prec now, hold(e_2, t_2), greet(max, john, e_2)]$

The DRP representing (1) is constructed by updating the DRP $\{\{\alpha\}, \emptyset\}$ with β.

In section 3.1, we argued in favour of using the defeasible LK Narration.

- **Narration**
  If the clause β is to be attached to the clause α, then normally, $Narration(\alpha, \beta)$ holds.

We now formalise this. Let $\langle \tau, \alpha, \beta \rangle$ be a function that updates the DRP τ with the DRS β via a discourse relation with α, where α is an open constituent of τ. Thus the output of this function is a new DRP τ', which is just like τ save that β is added to the first set and a discourse relation between α and β added to the second set. Narration is then represented in CE as a schema:

- **Narration**
  $\langle \tau, \alpha, \beta \rangle > Narration(\alpha, \beta)$

By the above assumptions (a) to (e) on the reader's KB, it contains α, β, and $\langle \alpha, \alpha, \beta \rangle$. Thus $Narration(\alpha, \beta)$ follows from the KB by Defeasible Modus Ponens. So the DRP representing (1) is (1′).

(1) Max stood up. John greeted him.

(1′) $\{\{\alpha, \beta\}, \{Narration(\alpha, \beta)\}\}$

Now the task is to infer the relations between the events. Let $me(\alpha')$ stand for *the main eventuality described in* α'. The formal definition of me is given in the appendix in a way that agrees with intuitions: For example, $me(\alpha)$ is the event e_1 of Max standing up, and $me(\gamma)$ is the *state* $CS(e)$ of Max having taken the bus, where γ is the logical form of (14).

(14) Max had taken the bus.

(γ) $[s,t][t \prec now, hold(s,t), s = CS(e), take(bus, max, e)]$

The Axiom on Narration introduced earlier is represented as follows:

- **Axiom for Narration**
  $\Box(Narration(\alpha,\beta) \rightarrow me(\alpha) \prec me(\beta))$

In words, $Narration(\alpha,\beta)$ entails that α's eventuality occurs before β's. By modal closure in CE, one concludes from (1′) that the standing up precedes the greeting, in agreement with intuitions.

The representation of (1) is context-specific. If further information is added to the KB, perhaps obtained from the preceding discourse context, then $Narration(\alpha,\beta)$ may not be inferred because of the nonmonotonicity of the inference.

Discourse structure is related to topic structure, as defined in Asher (in press). Intuitively, the topic of a segment is the overarching description of what the segment is about. For example, the topic of (1) could be John's introduction to Max. In the case of *Narration*, the law below states that $Narration(\alpha,\beta)$ holds only if α and β have a distinct, common (and perhaps implicit) topic γ:

- **Topic of Narration**
  $\Box(Narration(\alpha,\beta) \rightarrow (\exists\gamma)(topic(\alpha) = \gamma \wedge topic(\beta) = \gamma \wedge \gamma \neq \alpha \wedge \gamma \neq \beta))$

This rule is used to explain the incoherence of (17): as long as WK about cars breaking down and the sun setting is represented as intuitions would dictate, one cannot find a common distinct topic, and so *Narration* between the clauses can't be inferred.

(17) ?My car broke down. The sun set.

But no other relation can be inferred given the defeasible LK and WK available. And hence no discourse structure for (17) is constructed. Note that we can improve the coherence of (17) by supplying a topic in a third clause, e.g. (18):

(18) Then I knew I was in trouble.

7.2 Explanation and The Penguin Principle

Now consider text (2).

(2) Max fell. John pushed him.

The representation of the clauses are α and β; they are constructed in the same way as the clauses in (1), since they have similar syntax.

(α) $[e_1, t_1][t_1 \prec now, hold(e_1, t_1), fall(max, e_1)]$

(β) $[e_2, t_2][t_2 \prec now, hold(e_2, t_2), push(john, max, e_2)]$

We represent the Push Causal Law and the law that Causes Precede Effects in CE as:

- **Causal Law 2**
  $\langle \tau, \alpha, \beta \rangle \wedge fall(max, me(\alpha)) \wedge push(john, max, me(\beta)) > cause(me(\beta), me(\alpha))$
- **Causes Precede Effects**
  $\Box(cause(e_1, e_2) \rightarrow \neg e_2 \prec e_1)$

We also capture the LK that if β is to be attached to α, and moreover β's event caused α's then normally, β explains α:

- **Explanation**
  $\langle \tau, \alpha, \beta \rangle \wedge cause(me(\beta), me(\alpha)) > Explanation(\alpha, \beta)$

And finally, in accordance with Scha and Polanyi (1988), we assume that the same two clauses cannot be related by both *Explanation* and *Narration*:

- **Axiom on Explanation**
  $\Box(Explanation(\alpha, \beta) \rightarrow \neg Narration(\alpha, \beta))$

It should be stressed that Explanation is defeasible because there are exceptions to the rule, which we will discuss shortly. But first, we analyse text (2). The KB contains α, β and $\langle \alpha, \alpha, \beta \rangle$. Thus the antecedent to two defeasible laws are verified: Narration and the Push Causal Law. The result is a complex Penguin Principle; it is complex because the consequent of the two defeasible laws are not χ and $\neg\chi$, but instead the laws conflict in virtue of the Axiom on Narration and Causes Precede Effects. CE supports the more complex Penguin Principle:

- **The Complex Penguin Principle**
  $\Box(\phi \rightarrow \psi), \psi > \chi, \phi > \zeta, \Box(\chi \rightarrow \theta), \Box(\zeta \rightarrow \neg\theta), \phi \mathrel{|\!\!\approx} \zeta$

Therefore, there is a defeasible inference that the pushing caused the falling, as required.

Given the results of the Complex Penguin Principle, the antecedent of Explanation is verified. Moreover, the antecedent to Explanation entails that of the conflicting law Narration. So there is another Complex Penguin Principle, from which $Explanation(\alpha, \beta)$ is inferred. So ($2'$) is the representation of (2).

(2′) $\{\{\alpha, \beta\}, \{Explanation(\alpha, \beta)\}\}$

The second application of the Penguin Principle in the above used the results of the first, but in nonmonotonic reasoning one must be wary of dividing theories into 'subtheories' in this way. We show in the appendix that the predicates involved in the above deduction are sufficiently independent that in CE one can indeed divide the above into two applications of the Penguin Principle to yield inferences from the theory as a whole. We call this double application of the Penguin Principle where the second application uses the results of the first the *Cascaded Penguin Principle.*

By the openness constraints on DRPs defined earlier, the only available sentence for attachment if one were to add a sentence to (1) is *John greeted him*, whereas in (2), both sentences are available. So although the sentences in (1) and (2) have similar syntax and logical forms, the texts have very different discourse structures. The events they describe also have different causal structures. The difference in the natural interpretations of (1) and (2) arose essentially from causal knowledge.

Again, it must be stressed that these inferences could change if the reader's KB is different. The difference between (1) and (2) is not that (1) is always a narrative and (2) is always an explanation. Rather, *in the absence of information to the contrary*, (1) is a narrative and (2) is an explanation. In (Lascarides, Asher and Oberlander 1992), we demonstrate that this framework is sufficiently powerful to explain why the natural interpretation of (2) *in vacuo* is different from the interpretation of the same sentence pair in (19).

(19) John and Max were at the edge of a cliff. Max felt a sharp blow to the back of his neck. <u>He fell. John pushed him.</u> Max rolled over the edge of the cliff.

In (19), the falling and pushing are related by *Narration* rather than *Explanation.*

The relation *Explanation* determines the topic of the text segment as described in Asher (in press): the rule below states that the topic of an explanation is the clause that's explained:

- **Topic of Explanation**
  $\Box(Explanation(\alpha, \beta) \rightarrow topic(\beta) = \alpha)$

Support for this rule, and for the defeasible nature of Explanation, comes from explaining incoherence of texts like (20) (first cited in Caenepeel 1990):

(20) ?Everyone laughed. Fred told a joke.

In (20), there is likely to be a causal law: if laughing and telling a joke are connected, then normally the latter caused the former. Accordingly, the reader will infer this causal relation, so Explanation's antecedent is verified. But in this 'null' context, it is hard to see how laughing (on its own) could be sufficiently 'important' or unusual to be the topic of an explanatory text segment. Assuming this knowledge is formally represented (cf. Grosz and Sidner 1986), the reader will infer that laughing cannot be the topic, and by the contrapositive version of Topic of Explanation, $\neg Explanation(\alpha, \beta)$ will be inferred. This blocks the conclusion of the default law Explanation (note that if Explanation were made indefeasible, then an inconsistent KB would arise). Moreover, *Narration* can't be deduced either, since this is blocked by the causal relation already inferred. Hence no discourse structure for (20) is constructed.[7] So, although (20) is temporally coherent in that a temporal relation is inferred, it is discourse incoherent because no discourse relation can be inferred.

7.3 Elaboration and The Penguin Principle

Now consider text (5).

(5) The council built the bridge. The architect drew up the plans.

The Building Law below represents the knowledge that if building the bridge and drawing the plans are connected, then normally the latter is in the preparatory phase of the former; $Info(\alpha, \beta)$ is a gloss for the information that α's and β's main eventualities are respectively building the bridge and drawing the plans.

- **Building Law**
  $\langle \tau, \alpha, \beta \rangle \wedge Info(\alpha, \beta) > prep(me(\beta), me(\alpha))$

Elaboration, Axiom on Elaboration and Topic of Elaboration represent LK similar to that for *Explanation*:

- **Elaboration**
  $\langle \tau, \alpha, \beta \rangle \wedge prep(me(\beta), me(\alpha)) > Elaboration(\alpha, \beta)$
- **Axiom on Elaboration**
  $\Box(Elaboration(\alpha, \beta) \rightarrow \neg Narration(\alpha, \beta))$
- **Topic of Elaboration**
  $\Box(Elaboration(\alpha, \beta) \rightarrow topic(\beta) = \alpha)$

Assuming the KB is as usual, $prep(me(\alpha), me(\beta))$ and $Elaboration(\alpha, \beta)$ are inferred by the Cascaded Penguin Principle. The pairs of conflicting laws are: The Building Law and Narration and Elaboration and Narration.

7.4 Background and the Penguin Principle

States Overlap, which was introduced earlier, is formally represented as follows:

- **States Overlap**
  $\langle \tau, \alpha, \beta \rangle \wedge state(me(\beta)) > overlap(me(\alpha), me(\beta))$

[7] Replacing *Everyone laughed* with *Everyone laughed until their sides split* improves coherence, even though the causal relation is the same. One can envisage a purpose for the text that makes the latter clause a suitable topic, and worth explaining.

Background captures the intuition that if β is to be attached to α and the eventualities they describe overlap, then the discourse relation is *Background*, if things are normal, and the Axiom on Background prevents the text segment from being iconic:

- **Background**
  $\langle \tau, \alpha, \beta \rangle \wedge overlap(me(\alpha), me(\beta)) > Background(\alpha, \beta)$
- **Axiom on Background**
  $\Box(Background(\alpha, \beta) \rightarrow \neg me(\alpha) \prec me(\beta))$

The appropriate KB when interpreting (3) forms the premises of a Cascaded Penguin Principle: the pairs of conflicting laws are States Overlap and Narration and Background and Narration

(3) Max opened the door. The room was pitch dark.

Thus the event and state overlap and the clauses are related by *Background*. Now we turn to text (4).

(4) Max switched off the light. The room was pitch dark.

We have the following causal law, which reflects the knowledge that the room being dark and switching off the light, if connected, are normally such that the event causes the state.[8]

- **Dark Causal Law**
  $\langle \tau, \alpha, \beta \rangle \wedge switchoff(max, light, me(\alpha)) \wedge dark(room, me(\beta)) > cause(me(\alpha), me(\beta))$

Suppose the reader's KB is as usual. Then this KB verifies the antecedents of Narration, Dark Causal Law and States Overlap. Narration conflicts with States Overlap, which in turn conflicts with the causal law. Moreover, the causal law is more specific than States Overlap (for we assume $dark(room, s_2)$ entails $state(s_2)$ by the stative classification of the predicate $dark$), and States Overlap is more specific than Narration. So there appear to be two Penguin Principles, much like the Cascaded Penguin Principle. But the predicates in this example are not sufficiently independent that we can divide the theory into two subtheories, as is required in the Cascaded Penguin. We in fact cannot divide the theory up at all. We give the full details of what discourse relations are inferred in the appendix. The proof demonstrates that because the causal law is the most specific one, a causal relation between the event and state is inferred. We call this pattern of inference the Double Penguin Principle.

A constraint on the use of the discourse relation *Result* is defined as follows:

- **Result**
  $\langle \tau, \alpha, \beta \rangle \wedge cause(me(\alpha), me(\beta)) > Result(\alpha, \beta)$

[8] For the sake of simplicity we ignore the problem of inferring that the light is in the room.

Given the inference derived from the above Double Penguin Principle, the antecedents to Result and Narration are verified. Result and Narration don't conflict, so the consequents of both Narration and Result are inferred from the KB. Thus text (4) is a narrative, and moreover, the darkness is the *result* of switching off the light.

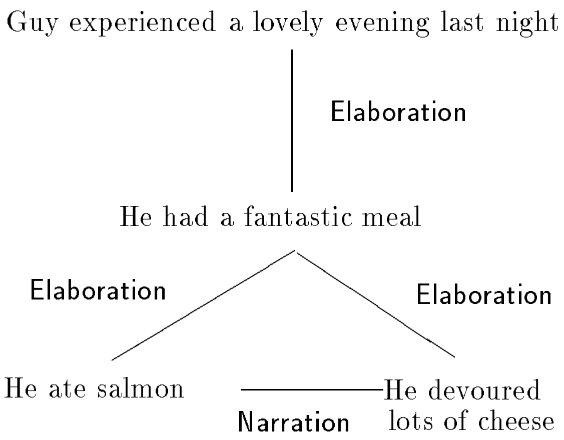
The Penguin Principle captures the following intuition concerning text processing: the reader never ignores information that is derivable from the text and relevant to calculating temporal structure and discourse structure. For example, in the case of (3) the Penguin Principle ensures the following information is not ignored: a stative expression was used instead of an event one.

7.5 Discourse Popping and The Nixon Diamond

Consider text (7).

(7) a. Guy experienced a lovely evening last night.
b. He had a fantastic meal.
c. He ate salmon.
d. He devoured lots of cheese.
e. He won a dancing competition.

Let the logical forms of the respective clauses (7a-e) be α, β, γ, δ and ϵ. The discourse structure for (7a-d) involves Cascaded Penguin Principles and Defeasible Modus Ponens as before. Use is made of the defeasible knowledge connecting eating cheese and salmon with having a meal and experiencing a lovely evening.



We study the attachment of ϵ to the preceding text in detail. The open clauses are δ, β and α. So by the assumptions on the contents of the KB, the reader believes $\langle\tau,\alpha,\epsilon\rangle$, $\langle\tau,\beta,\epsilon\rangle$ and $\langle\tau,\delta,\epsilon\rangle$. $\langle\tau,\delta,\epsilon\rangle$ verifies the antecedent to Narration, but intuitively δ is not related to ϵ at all. We must therefore block the inference to Narration's consequent.

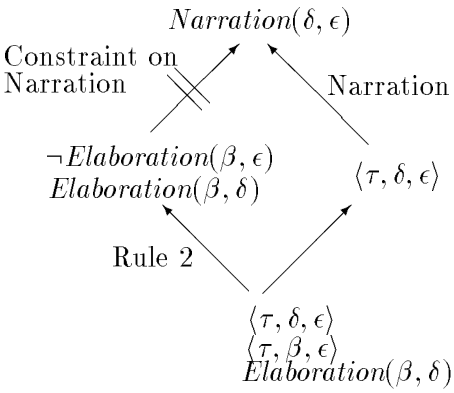
To do this we use the intuition that since δ elaborates β, ϵ can normally form a narrative with δ only if ϵ also elaborates β. This is captured by the following:

- **Constraint on Narration**
  $Elaboration(\beta, \delta) \wedge \neg Elaboration(\beta, \epsilon) > \neg Narration(\delta, \epsilon)$

Moreover, its antecedent is verified by the KB. For $Elaboration(\beta, \delta)$ is already in the KB; and $\neg Elaboration(\beta, \epsilon)$ is infered from Rule 2, which is obtained from Narration and the contrapositive version of Axiom on Elaboration, via Closure on the Right (β and ϵ have been substituted in the schema).

- **Rule 2**
  $\langle \tau, \beta, \epsilon \rangle > \neg Elaboration(\beta, \epsilon)$

The result is a 'Nixon Polygon'. There is *irresolvable conflict* between Narration on the one hand, and Rule 2 and the Constraint on Narration on the other because their antecedents are not logically related:



We show in the appendix that similarly to the Nixon Diamond, a Nixon Polygon yields no conclusions: neither $Narration(\delta, \epsilon)$ nor $\neg Narration(\delta, \epsilon)$ are inferred from the above.

We assume that any KB that verifies $\langle \tau, \delta, \epsilon \rangle$ but *fails* to verify any discourse relation between δ and ϵ is inconsistent. Therefore, the above assumption that $\langle \tau, \delta, \epsilon \rangle$ is part of the KB must be dropped. The next KB to be considered is the same as the original save that it *doesn't* contained $\langle \tau, \delta, \epsilon \rangle$. So ϵ must be attached to α or β, but not δ. Using Cascaded Penguin Principles and Defeasible Modus Ponens, one infers $Elaboration(\alpha, \epsilon)$ and $Narration(\beta, \epsilon)$, in agreement with intuitions.

The above shows that the Nixon Diamond provides the key to discourse popping; ϵ is attached to a clause that dominates the previous sentence δ because a Nixon Diamond meant that ϵ couldn't be attached to δ itself.

7.6 Dudley Doorite and Text Ambiguity

We earlier indicated that interpreting text (9) requires Closure on the Right and Dudley Doorite.

(9) The bimetallic strip changed shape. The temperature fell.

We used these inferences to motivate the following causal law:

- **Change Causal Law**
  $\langle \tau, \alpha, \beta \rangle \wedge change(strip, me(\alpha)) \wedge fall(temperature, me(\beta)) >$
  $cause(me(\alpha), me(\beta)) \vee cause(me(\beta), me(\alpha))$

Now suppose we assume that the sentences in (9) are *not* dominated by a common topic. This conjecture is supported by standard heuristics for defining constraints on possible topics, which entail that sentences under a common topic should share actors and/or patients (cf. Grosz & Sidner 1986, Asher in press). Then, by the contrapositive version of Common Topic for Narration, the relevant KB in analysing (9) verifies $\neg Narration(\alpha, \beta)$. So although the antecedent to Narration is satisfied in analysing (9), $Narration(\alpha, \beta)$ cannot be inferred.

Closure on the Right and Dudley Doorite yield the following from Result and Explanation:

Closure on the Right on Result and Explanation
(i) $\langle \tau, \alpha, \beta \rangle \wedge cause(me(\alpha), me(\beta)) > Result(\alpha, \beta) \vee Explanation(\alpha, \beta)$
(ii) $\langle \tau, \alpha, \beta \rangle \wedge cause(me(\beta), me(\alpha)) > Result(\alpha, \beta) \vee Explanation(\alpha, \beta)$

Dudley Doorite on (i) and (ii)
(iii) $\langle \tau, \alpha, \beta \rangle \wedge (cause(me(\alpha), me(\beta)) \vee cause(me(\beta), me(\alpha))) >$
$(Result(\alpha, \beta) \vee Explanation(\alpha, \beta))$

The antecedent of Change Causal Law is satisfied by the KB. So by Defeasible Modus Ponens $cause(me(\alpha), me(\beta)) \vee cause(me(\beta), me(\alpha))$ is inferred. By Defeasible Modus Ponens again on rule (iii), (21) is inferred.

(21) $Explanation(\alpha, \beta) \vee Result(\alpha, \beta)$

Thus the discourse relation between the sentences in (9) is ambiguous.[9]

8 The Pluperfect

We have so far examined how to infer discourse relations in simple past tensed texts. We now extend the theory to cover texts that feature the pluperfect. As we have mentioned, a pluperfect clause like (14) is represented as (14′).

(14) John had taken the bus.

[9] Note that this is an informative ambiguity in that $Explanation(\alpha, \beta) \vee Result(\alpha, \beta)$ is not a tautology.

$(14')$ $[s,t][t \prec now, hold(s,t), s = CS(e), take(john, bus, e)]$

The following axioms relate $CS(e)$ and e:

- **The Relation between $CS(e)$ and e**
  $\Box(\forall et)(hold(CS(e),t) \rightarrow (\exists t')(t' \prec t \wedge hold(e,t')))$
  $\Box(\forall et)(hold(e,t) \rightarrow (\exists t')(t \prec t' \wedge hold(CS(e),t')))$

Given the above axioms, pluperfect sentences invoke equivalent temporal structures to the corresponding simple past tensed ones (this equivalence assumes time is dense). The only difference between (14′) and the logical form (22′) of (22) lies in which eventuality is accessible for future anaphoric reference.

(22) Max took the bus.

$(22')$ $[e,t][t \prec now, hold(e,t), take(john, bus, e)]$

This is unlike previous representations of the pluperfect, which invoke a tripartite relation between event time, reference time and speech time (Kamp & Rohrer (1983), Partee (1984), Hinrichs (1986), Dowty (1986)). In these theories, the simple past and pluperfect tenses differ at the sentential level because the reference time is identified with the event time in the simple past and is after the event time in the pluperfect; thus the temporal structures are different.

But even though there is sentential equivalence, the pluperfect will play a distinct discourse role, as explained below. Because discourse structure doesn't arise solely from the logical form of clauses, the theory is rich enough to distinguish between the truth conditions of sentences and their discourse roles, thus allowing us to contrast the simple past and pluperfect in this way. Exploring this discourse level strategy for analysing the pluperfect is motivated by the fact that pluperfect sentences are understandable only in the context of preceding simple past tensed discourse.

To explore what role the pluperfect plays in discourse, we first consider the following texts.

(23) ?Max poured a cup of coffee. He had entered the room.

(24) Max entered the room. He poured a cup of coffee.

Intuitively, the temporal order between entering the room and pouring a cup of coffee are the same in (23) and (24). So why is (23) awkward but (24) acceptable? Intuitively, we can find a connection between the events in (24) that is compatible with the tenses used; the connection is that Max entered the room so that he could pour himself a cup of coffee. On the other hand, this connection cannot be compatible with the tenses used in (23), since (23) is awkward. Theories that analyse the distinction between the simple past and pluperfect purely in terms of different relations between reference times and event times, rather than in terms of event-connections, fail to explain why (30) is acceptable but (23) is awkward.

Intuitively, a pluperfect clause β can be connected to a simple past tense clause α if β elaborates α (e.g. text (25)) or explains α (e.g. (26)).

(25) Max arrived late for work. He had taken the bus.

(26) The council built the bridge. The architect had drawn up the plans.

Indeed, it suffices for β to be *part of* an elaboration or explanation of α, for we can improve text (23) by adding to it a further clause that, together with the pluperfect clause, *explains* Max pouring the coffee (e.g. text (27), first cited in Caenepeel 1991):

(27) Max poured himself a cup of coffee. He had entered the room feeling depressed, but now felt much better.

Furthermore, pluperfect and simple past clauses can form a 'parallel' under a common topic, as in text (28) (cf. Asher in press), and contrast each other, as in (29).

(28) Max was wrong. He had been wrong. He will always be wrong.

(29) Max got the answer wrong. But John had got the the answer right.

Both the contrast and parallels must be such that the event in the pluperfect happens before that in the simple past.

This seems to exhaust the possible connections between the simple past event and the pluperfect one. This is why (23) is awkward, for it is difficult to think of the circumstances under which the entering the room explains or elaborates pouring the coffee, and they clearly cannot form parallel or contrast. Of course, many more connections are allowed between two events described in the simple past. The task now is to formalise this intuition that the discourse role of the pluperfect is to restrict the kinds of discourse connections that can hold.

Let $C_{pp}(\alpha, \beta)$ mean that α and β are connected by the kind of discourse relation allowed between simple pasts and pluperfects: that is, β is (part of) an elaboration or explanation of α, or β and α form parallel or contrast, with the *event* described in β preceding that described in α. Then the discourse role of the pluperfect is captured as follows:

- **Connections When Changing Tense** (CCT)
  $\Box(\langle \tau, \alpha, \beta \rangle \wedge sp(\alpha) \wedge pp(\beta) \rightarrow C_{pp}(\alpha, \beta))$

It's important to stress that $C_{pp}(\alpha, \beta)$ *doesn't* necessarily follow if β is simple past rather than pluperfect. For simple past tensed texts, more connections are permissable, e.g., $Narration(\alpha, \beta)$.

Given the above rule one can explain the awkwardness of (23). If the WK is stated as intuitions would dictate, then the KB in analysing (23) would be one where the clauses described in (23) are not related by C_{pp}. So by the contrapositive version of CCT and the fact that information states incorporate modal closure, the sentences are not related by a discourse relation; for $\langle \tau, \alpha, \beta \rangle$ must be false. This makes (23) incoherent. Analysing (27) would involve relating whole discourse segments to a clause by a discourse relation; that is the second and third clauses *together* must stand in the relation *Explanation* to the first. Providing such analysis is beyond our scope here since we have concentrated on relations between single clauses (but see Lascarides and Asher (in preparation) for a detailed discussion of such cases).

CCT captures the relevant distinction between the use of the simple past tense and the pluperfect as illustrated in (30) and (31).

(30) Max slipped. He had spilt a bucket of water.

(31) Max slipped. He spilt a bucket of water.

Intuitively, the discourse relation in (30) is *Explanation* and in (31) it is *Narration*. This difference must arise solely in virtue of the pluperfect. The analysis of (31) is exactly like that of (1): by Defeasible Modus Ponens on Narration, it's narrative. Now consider (30). The relevant KB satisfies the antecedents of Narration, States Overlap and the indefeasible law CCT. CCT entails $C_{pp}(\alpha, \beta)$, where α and β are respectively the logical forms of the clauses. Now if spilling the water and slipping are connected so that either the spilling explains, elaborates, parallels or contrasts the slipping, then normally the spilling explains the slipping. This intuition is reflected in the following law, where $Info(\alpha, \beta)$ is the gloss for "α describes Max slipping and β describes Max spilling a bucket of water".

- **Slipping Law**
  $\langle \tau, \alpha, \beta \rangle \wedge C_{pp}(\alpha, \beta) \wedge Info(\alpha, \beta) > Explanation(\beta, \alpha)$

Thus the KB satisfies the antecedents to the above law as well. Note that the Slipping Law is different from the causal laws considered so far, because $C_{pp}(\alpha, \beta)$ is in the antecedent. This added conjunct is crucial to the plausibility of the above law concerning spilling and slipping. For intuitively, there is *no* WK that normally the spilling caused the slipping if the choice of connection between these events encompasses the full range of possibilities, including that the slipping caused the spilling. It must be also stressed that this law plays no part in the analysis of (31); the relevant KB doesn't verify its antecedent, because $C_{pp}(\alpha, \beta)$ is not in the KB.

We must investigate what follows from Narration, States Overlap and the Slipping Law in the interpretation of (30). The Slipping Law and States Overlap don't conflict, but they each conflict with Narration and are more specific than it. By the Double Penguin Principle, $Explanation(\alpha, \beta)$ and $overlap(me(\alpha), me(\beta))$ are inferred; thus the antecedent to Background is verified and $Background(\alpha, \beta)$ is inferred. So, spilling the water explains *why* Max slipped, and the consequences of spilling the water are are still in force when Max slips.

These texts show that although simple past tensed sentences are equivalent to the corresponding pluperfects they play distinct roles in discourse. Our formalism reflects the intuition that the pluperfect acts as a syntactic discourse marker to indicate that only a restricted set of discourse relations is possible, thus yielding different inferences about discourse structure. We have used indefeasible and defeasible inference on LK to explain this and not formal devices like reference times.

We have investigated how the pluperfect sentences are related to simple past tensed ones. We now evaluate how pluperfect sentences are related to each other. Consider (32).

(32) a. Max arrived late for work.

b. He had taken the bus.

c. He had totalled his car.

d. His boss summoned him to his office.

The theory presented here is rich enough to reflect the intuition that each pluperfect sentence in (32) explains the previous pluperfect sentence. Let the logical forms of (32a) to (32d) be respectively α, β, γ and δ. The discourse relation between α and β is worked out an analogous way to that in (30): $Background(\alpha, \beta)$ and $Explanation(\alpha, \beta)$ are inferred. α and β are both open clauses to γ. Again, by a similar inference to that for text (30) $Background(\alpha, \gamma)$ and $Explanation(\alpha, \gamma)$ are inferred.

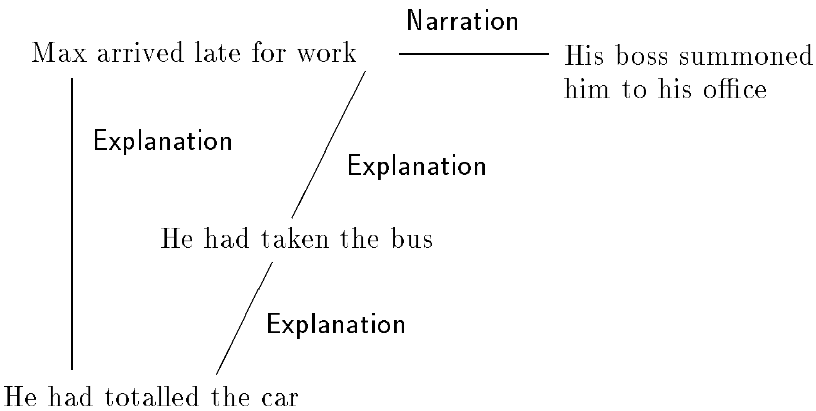
In connecting β to γ, the KB satisfies the antecedent of States Overlap and Narration. There is a also a causal law relating totalling the car and taking the bus. By the Penguin Principle, the consequents of States Overlap and the Causal Law are inferred, and subsequently, we infer $Background(\beta, \gamma)$ and $Explanation(\beta, \gamma)$.

The open clauses to δ are α, β and γ. A Nixon Polygon occurs when attempting to attach δ to γ and δ to β in a similar way to that which occurred in (7) described earlier. This time, we use a constraint on Narration defined in terms of *Explanation* rather than *Elaboration*:

- **Constraint 2 on Narration**
  $Explanation(\beta, \gamma) \wedge \neg Explanation(\beta, \delta) > \neg Narration(\gamma, \delta)$

Thus δ must attach to α. By Defeasible Modus Ponens on Narration, $Narration(\alpha, \delta)$ is inferred. Thus the DRP representing (32) is (32′).

(32′) $\{\{\alpha, \beta, \gamma, \delta\},\ \{Explanation(\alpha, \beta), Background(\alpha, \beta)$
$Explanation(\alpha, \gamma), Background(\alpha, \gamma)$
$Explanation(\beta, \gamma), Background(\beta, \gamma)$
$Narration(\alpha, \delta)\}\}$



The pluperfect sentences in (32) each create further subordinations in the discourse structure. But pluperfects don't necessarily do this. Consider (33).

(33) a. Alexis was a really good girl by the time she went to bed yesterday.

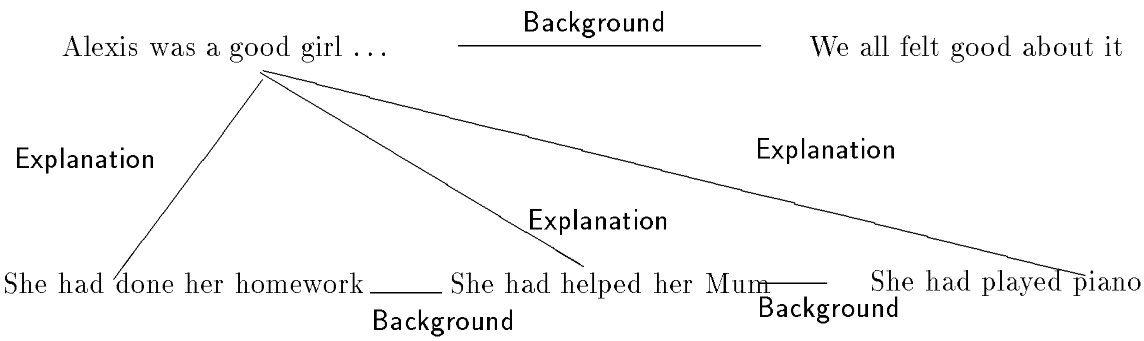
b. She had done her homework.

c. She had helped her Mum with housework.

d. She had played piano.

e. We all felt good about it.

Suppose the logical forms of (33a) to (33e) are respectively α to ϵ. By a similar argument to the one above, common sense inference yields $Explanation(\alpha, \beta)$, $Explanation(\alpha, \gamma)$ and $Explanation(\alpha, \delta)$. World knowledge does not yield any preference as to the connections between the *event* in γ and the *event* in the open clause β. So β and γ satisfy the antecedents of States Overlap and Narration, and so a Cascaded Penguin Principle yields $Background(\beta, \gamma)$. Similarly, we infer $Background(\gamma, \delta)$. And as in (32), a Nixon Diamond gives rise to discourse 'popping' for ϵ, and in fact we obtain $Background(\alpha, \epsilon)$.



The fact that (33b-d) form background to each other entails nothing about how the events they describe are temporally ordered, for we cannot infer from *Background* where the the consequent states start relative to each other, and the corresponding events are placed at the start of these consequent states. This is in agreement with the intuitive interpretation of (33), in contrast to the order of events inferred from Narration when analysing the corresponding simple past tensed text (34).

(34) Alexis did her homework. She helped her Mum with housework. She played piano.

The events described by a sequence of pluperfect clauses will occur in their textual order only if that order is inferrable from WK or other temporal information present in the text (cf. (35)).

(35) Max arrived at the summit at midday. He had got up at 5:30am, had prepared his lunch, had chosen his route, and had passed based camp before 7am.

This is in agreement with the claim made in Kamp (1991), that the temporal order of events described by a sequence of pluperfect clauses depends on the discourse type.

9 Conclusion

We have proposed a framework for calculating the structure of text, where the contributions made by syntax, semantic content, causal knowledge and linguistic knowledge are all explored within a single logic. The distinct discourse relations that are prevalent in texts with similar syntax arose from defeasible laws, by using intuitively compelling patterns of defeasible entailment.

We argued that a full account of the simple past and pluperfect tenses exemplify the inference patterns of Defeasible Modus Ponens, The Penguin Principle, the Nixon Diamond, Closure on the Right, and Dudley Doorite. Closure on the Right and Dudley Doorite provided the key to discourse structure ambiguity. The Nixon Diamond provided the key to 'popping' from subordinate discourse structure. And the Penguin Principle captured the intuition that the reader never ignores information relevant to calculating discourse structure.

We examined the effects of WK and LK on interpretation in a logical context. Consequently, the traditional DRT syntactic-based account for calculating relations between reference times in order to determine temporal structure was not needed. The distinct discourse roles of the simple past and the pluperfect were brought out through a mixture of defeasible and indefeasible reasoning, rather than by different relations between reference times determined in syntax.

Appendix

The Language

The basic building blocks of DRPs are DRSs and discourse relations between them. In order to simplify the proofs in this appendix, we make some simplifying assumptions about the language for reasoning about discourse structure. In particular, the language in this appendix will not make full use of the truth definitions of DRSs. We are able to simplify things in this way because the only semantic information that's *part* of a DRS and relevant to constructing discourse structure—such as the information that the event described is Max falling—can be represented as a predicate which takes the DRS as an argument. In other words, for simplicity, we re-write $fall(max, me(\alpha))$ as $Maxfall(\alpha)$. The second simplification concerns the function me (standing for "main eventuality"). In the text, me is a partial function me :DRS $\rightarrow U_{\mathrm{DRS}}$, where U_α is the universe of the DRS α and $me(\alpha)$ is the main eventuality discourse referent as defined in (Asher in press). But we can for the sake of reasoning about DRP construction rephrase the function me and relations involving main events of DRSs as relations on the DRSs themselves. These two moves make the proof theory simpler. In particular our language will remain that of propositional logic. We call this propositional language for reasoning about DRP construction L_{DRP}.

We can essentially treat DRSs as atomic propositions in L_{DRP}. But this simplification comes with a cost. L_{DRP} is capable of expressing the fact that a DRS is true, but this assignment is independent of the assignment of truth values to DRSs given in the standard truth definition for DRSs below. If we wish, we may restrict our attention to those L_{DRP} models in which the assignment of truth values to DRSs and to predicates involving DRSs reflects the values assigned in the standard interpretation of DRSs. But we leave open the question of how the proofs in this appendix must be extended so as to apply to this class of models.

The Standard Truth Definition for DRSs

We give here the semantics of atomic DRSs, which are the only kind used in this paper. A full truth definition for DRT can be found in Kamp and Reyle (in press).

- If ψ is an n-ary DRS predicate and $x_1, \ldots, x_n$ are discourse references, then $\psi(x_1, \ldots, x_n)$ is an atomic condition.
- If x_1 and x_2 are discourse references (of any kind), then $x_1 = x_2$ is an atomic condition.
- A DRS is a pair $\langle U, Con\rangle$, where U is a set of discourse referents, and Con is a set of conditions.

The semantics of DRSs are defined with respect to a model $\langle W, T, D, E, [\![\,]\!]\rangle$, where W is a non-empty set of possible worlds; T is a non-empty set of time points; D is a function from W into a family of non-empty sets ($D(w)$ is the domain of individuals in w); E is a function from W into a family of non-empty sets ($E(w)$ is the domain of eventualities in

w); and $[\![]\!]$ is an interpretation function that assigns to DRS predicates functions from W to $\wp(\cup_{n\in\omega}(D \cup E)^n \cup T)$.

A *proper embedding* of a DRS K in a model M is defined in terms of a function from discourse referents into objects O_M in M ($O_M = D_M \cup E_M \cup T_M$). Any such function from discourses references into O_M is an embedding function.

Define g to *extend* an embedding function f to an embedding of K in M at w (written $g \supseteq_K f$) just in case $Dom(g) = Dom(f) \cup U_K$.

Define an *external anchor* A for K in M to be a partial function from U_K into $\cup_{w\in W} D_w \cup T$, such that if *now* occurs in U_K then $A(now)$ is the utterance time of the discourse, and if i occurs in U_K, then $A(i)$ is the speaker of the discourse.

Now we define a *proper embedding* f of K in M at w (written $[f, K]^M w$) with respect to a possibly empty external anchor A for K in M and we define satisfaction of a condition in M relative to an embedding function f for the DRS in which the conditions occur at w (written $M \models_{w,f}$).

(i) If ψ is an atomic condition of the form $\phi(x_1, \ldots, x_n)$, then $M \models_{w,f} \psi$ iff $\langle f(x_1), \ldots, f(x_n)\rangle \in [\![\phi]\!](w)$.

(ii) If ψ is an atomic condition of the form $x_1 = x_2$, then $M \models_{w,f} \psi$ iff $f(x_1) = f(x_2)$.

(iii) If A is an external anchor for K in M, then $[f, K]^M_{w,g}$ iff (i) $f \supseteq_K g$; (ii) $A \subseteq f$; (iii) $\forall\theta \in Con_K$, $M \models_{w,f} \theta$.

Definition of Main Eventuality

To define $me(\alpha)$, we assume the theory is expressed in DRT with the syntax analysed in GPSG. Under these circumstances, the definition is as follows: For a non-conjunctive clause α, $me(\alpha)$ is the eventuality discourse entity introduced by the verb phrase that heads the clause.

The Language L_{DRP}

The language of L_{DRP} is built recursively in the usual way:

- DRS names α, β, γ are well formed expressions of L_{DRP}, and DRP names τ and τ' are well-formed expressions of L_{DRP}.
- If α is a DRS name or DRP name, then $\downarrow \alpha$ is a well-formed formula (WFF) of L_{DRP}.
- If $\alpha_1, \ldots, \alpha_n$ are DRS names, and p_n is an n-place predicate on DRS names, then $p_n(\alpha_1, \ldots, \alpha_n)$ is an atomic WFF.
- If α and β are WFF, then $\alpha \wedge \beta$, $\alpha \vee \beta$, $\neg\alpha$, $\alpha \rightarrow \beta$, $\alpha > \beta$ and $\Box\alpha$ are WFF.
- If τ is a DRP name and α and β are DRS names, then $\langle\tau, \alpha, \beta\rangle$ is a WFF.
- If ϕ, ψ and χ are WFF, then $(\phi, \psi) > \chi$ is a WFF.

The Truth Definition for L_{DRP}

The well-formed expressions of the language are defined with respect to a model $\langle W, R, *, f\rangle$ for the language, where f assigns DRS and DRP names functions from W to propositions (which in turn are function from W to $\{0,1\}$), and assigns predicate constants functions from W to suitable sets of propositions.

(a) Where β is either a name constant or a predicate constant, $[\![\beta]\!]^M(w) = f(\beta)(w)$.
Where β is a DRS or DRP name, $[\![\downarrow\beta]\!]^M(w) = (f(\beta)(w))(w)$.

(b) Where β is an atomic wff $p^n(d_1,\ldots,d_n)$,
$[\![\beta]\!]^M(w) = 1$ iff $\langle[\![d_1]\!]^M(w),\ldots,[\![d_n]\!]^M(w)\rangle$ belongs to $[\![p^n]\!]^M(w)$.

(c) $[\![\neg A]\!]^M(w) = 1$ iff $[\![A]\!]^M(w) = 0$.

(d) $[\![A \wedge B]\!]^M(w) = 1$ iff $[\![A]\!]^M(w) = 1$ and $[\![B]\!]^M(w) = 1$.
$\rightarrow$ and$\vee$ are defined in terms of $\wedge$ and $\neg$ in the usual way.

(e) $[\![\phi > \psi]\!]^M(w) = 1$ iff $*(w,[\![\phi]\!]^M) \subseteq [\![\psi]\!]^M$

(f) $[\![\Box\phi]\!]^M(w) = 1$ iff for all w' such that $R(w,w')$, $[\![\phi]\!]^M(w') = 1$

(g) $[\![(\phi,\psi) > \chi]\!]^M(w) = 1$ iff $*(w,[\![\phi]\!]) \cap *(w,[\![\psi]\!]) \subseteq [\![\chi]\!]$

Constraints on the Model

We assume that R is the S5 accessibility relation on worlds. Furthermore, the function $*$ must satisfy the following constraints for the model to be admissable:

- **Facticity**
  $*(w,p) \subseteq p$
- **Specificity**
  If $p \subseteq q$, $*(w,p) \cap *(w,q) = \emptyset$, then $*(w,q) \cap p = \emptyset$
- **Dudley Doorite**
  $*(w,p \cup q) \subseteq *(w,p) \cup *(w,q)$

Facticity captures the intuition that whatever the properties of a normal bird, one is that he's a bird. Specificity captures the intuition that normal birds aren't penguins. Dudley Doorite captures the intuition that if quakers are pacifists and republicans are pacifists, then quakers and republicans are pacifists.[10] We will henceforth restrict ourselves to admissable models.

Monotonic Proof in CE

The monotonic theory of CE in (Asher and Morreau 1991) is to be strengthened by the following axiom schemata which reflect S5 'modal closure' instead of just logical closure.

[10] In fact, Specificity follows from Dudley Doorite.

(A1) ϕ where ϕ is a tautology.

(A2) $\phi > \phi$

(A3) $(\phi > \psi \wedge \zeta > \psi) \rightarrow ((\phi \vee \zeta) > \psi)$

(A4) $((\phi, \psi) > \perp \wedge \Box(\phi \rightarrow \psi)) \rightarrow \psi > \neg\phi$

(A5) $(\phi > \psi) \rightarrow ((\phi, \psi) > \psi)$

(A6) $((\phi, \psi) > \chi) \leftrightarrow ((\psi, \phi) > \chi)$

(A7) $(\phi > \chi \wedge \psi > \chi) \rightarrow (\phi, \psi) > \chi$

(A8) $(\phi, \phi) > \psi \rightarrow \phi > \psi$

(A9) $\Box(\phi \leftrightarrow \psi) \rightarrow ((\phi > \zeta) \leftrightarrow (\psi > \zeta))$

(A10) $\Box(\phi \rightarrow \psi) \rightarrow ((\zeta > \phi) \rightarrow (\zeta > \psi))$

(A11) $\Box\phi \rightarrow \phi$

(A12) $\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$

(A13) $\Diamond\phi \rightarrow \Box\Diamond\phi$.

(R) $((\phi \rightarrow \psi) \wedge \phi) \rightarrow \psi$

(R1) If $\vdash \phi$, then $\vdash \Box\phi$

(R2) $\Box((\phi_1 \wedge \ldots \phi_n) \rightarrow \phi) \rightarrow ((\psi > \phi_1 \wedge \ldots \psi > \phi_n) \rightarrow \psi > \phi)$

Let the above axioms and rules be called T_0. We will write $\vdash_{T_0}$ for the derivability notion, and we define the L_{DRP} consequence to be the following:

- L_{DRP} **Consequence**
  $\Gamma \models_{L_{\mathrm{DRP}}} \phi$ iff for every admissible L_{DRP} model M and for every world w in W_M if every element of Γ is true at w in M then ϕ is true at w in M.

Fact:
$\Gamma \vdash_{T_0} \phi \Rightarrow \Gamma \models_{L_{\mathrm{DRP}}} \phi$

To prove completeness, we construct a canonical model in the usual fashion for propositional modal logic, except that we add a construction for $*$. Each world in the set of worlds ☺ in the canonical model is defined to be a maximal consistent set of sentences, and we may for each such set Γ and each formula ψ define:

$\Gamma^+_\psi = \{\phi : (\psi > \phi) \in \Gamma\}$
$\Gamma^-_\psi = \{\phi : \neg(\psi > \phi) \in \Gamma\}$

$*(\Gamma, [\![\phi]\!]) = \{\Gamma' \in W : \Gamma^+_\phi \subseteq \Gamma'\}$.

This construction suffices to yield a canonical model such that one can then prove:

Fact:
$\Gamma \models_{L_{\mathrm{DRP}}} \phi \Rightarrow \Gamma \vdash_{T_0} \phi$

We mention one other result about propositional CE, which one can prove using the filtration method (cf. Chellas 1980):

Fact: T_0 has the finite model property, and so is decidable.

The Nonmonotonic Part of CE

Just as in (Asher and Morreau 1991), we define CE entailment, written $\approx\!\!\!\mid$, in terms of the function $*$ on information states, the update function $+$, and the normalisation function N. These three functions are defined as follows:

Definition of $*$ on information states: $*(s,p) = \cup_{w \in s} * (w,p)$
Definition of $+$: $s + \Gamma = \{w \in s : w \models \Gamma\}$
Definition of Normalisation:

$$\begin{array}{rcl} N * (s,p) & = & \{w \in s : w \notin p \setminus *(s,p)\} \text{ if } *(s,p) \cap s \neq \emptyset \\ & = & s \text{ otherwise} \end{array}$$

The definition of $\approx\!\!\!\mid$ uses the notion of *normalisation chain* derived from N as follows:

Definition of $Ant(\Gamma)$: Where Γ is a set of sentences of L_{DRP} we define $Ant(\Gamma)$ to be $\{\phi : \phi > \psi$ occurs positively as a subset of a formula of $\Gamma\}$
$Cons(\Gamma)$—the set of consequents of conditionals in Γ—is defined analogously to $Ant(\Gamma)$.

Definition of Γ-normalisation chain: For $\Gamma \subseteq L_{\mathrm{DRP}}$, the Γ-*normalisation chain* with respect to an enumeration μ of $Ant(\Gamma)$ is defined to be the following sequence:
$N^0_\nu = s$
$N^{\alpha+1}_\nu = N(N^\alpha_\nu(\phi_i))$, where $\alpha = \lambda + n + 1$, and $\nu(\phi_i) = n + 1$
$N^\lambda_\nu = \cap_{\mu \in \lambda} N^\mu_\nu$

Definition of $\approx\!\!\!\mid$: $\Gamma \approx\!\!\!\mid \phi$ iff for any Γ-normalisation chain C beginning from ☺$+\Gamma$, $C^* \models \phi$, where C^* is the fixpoint of C, where ☺ is the set of worlds constructed in the canonical model.

When Γ is finite note that $Ant(\Gamma)$ is finite. And when $Ant(\Gamma)$ is finite and Γ has no embedded $>$-formulae within $>$-formulae, it is easy to see that all normalisation chains defined with respect to $Ant(\Gamma)$ reach a fixed point after finitely many normalisations. Together with the fact that T_0 has the finite model property, we can show the following:

Fact: $\approx\!\!\!\mid$ is decidable.

We will be working in the canonical model ☺. So in particular we will often define information states for various theories in terms of the operations of (Asher & Morreau 1991) — e.g. the information state S corresponding to a theory T will be defined as $S =$☺$+T$. The following definitions will also be useful.

Definition: Let T be a theory. $Formula(T) = \{\phi : \phi \in Ant(T) \vee \phi \in Cons(T) \vee T \models \phi\}$.

Definition: Let T be a theory. Then ϕ, ψ are T-independent iff any boolean combination of ϕ and ψ is consistent with T.

Definition: Two sets of formulae X, Y are T-independent iff $\forall\phi \in X$, $\forall\psi \in Y$, ϕ, ψ are T-independent.

Fact: Let T be any theory, and let $T_1, \ldots, T_n$ be subtheories such that every $Formula(T_i)$ and $Formula(T_j)$, $i \neq j$ are T-independent. Then $Formula(T_i) \cap Formula(T_j) = \emptyset$.

Definition: A T survivor world w is a world that survives on every normalisation chain defined with respect to $Ant(T)$ and which verifies T in ☺.

Definition: A default theory is a set of $>$ conditionals.

Definition: $* \upharpoonright (T, w)$ is the $*$ function restricted to $Ant(T)$ at w.

Definition: $Th(w \upharpoonright L_T) = \{\phi \in L_{\mathrm{DRP}} \cap L_T : w \models \phi\}$

We will use the term *theory* here in a colloquial sense: we will not assume as is standard in logic that a theory is closed under some notion of logical consequence. Rather, our theories are collections of sentences that constitute the reader's knowledge, as described in the paper. We now prove the lemmas that enable us to divide these theories into subtheories in CE under special circumstances. These will be used to model the Cascaded Penguin Principle.

Irrelevance Lemma
Suppose T is a theory and $T_1, \ldots, T_n$, $T_i \subseteq T$ are default theories such that $\cup T_i = T$. If $Formula(T_i)$ and $Formula(T_j)$ for $i \neq j$ are T-independent, then for any T_i and T_i survivor world w_i, if each T_j for $j = 1, \ldots n$ has a survivor world then there is a T survivor world w_1 such that $* \upharpoonright (T_i, w_1) = * \upharpoonright (T_i, w_i)$ and $Th(w_1 \upharpoonright Formula(T_i)) = Th(w_i \upharpoonright Formula(T_i))$.

Proof
Suppose $T_1, \ldots, T_n$ are as above and suppose all the individual T_j have survivor worlds. Let w_0 be a T_i survivor world. By assumption T_1 has a survivor world w_1. By independence, $\exists w_1'$ such that w_1' is a survivor world for T_1 and $Th(w_1', \upharpoonright Formula(T_i)) = Th(w_0, \upharpoonright Formula(T_i))$. Furthermore, by the assumption of independence, $Ant(T_1) \cap Ant(T_i) = \emptyset$. So we may choose w_1' such that for each $p \in Ant(T_i)$, $*(w_1', p) = *(w_0, p)$. Since by hypothesis each T_j is independent from T_i, we can repeat this for $T_2, \ldots, T_n$ and so construct a T survivor world. □

Strengthened Irrelevance Lemma
Suppose T is a theory and $T_1, \ldots, T_n$, $T_i \subseteq T$ are default theories such that $\cup T_i = T$. If $Formula(T_i) - \Delta$ and $Formula(T_j) - \Delta$ for $i \neq j$ are T-independent for some $\Delta \subseteq L_{T_k} \cap \ldots \cap L_{T_m}$ and there are survivor worlds $w_k, \ldots, w_\mu$ for $T_k, \ldots, T_m$ respectively such that for each $\delta \in \Delta$ and each w, $w' \in \{w_k, \ldots, w_\mu\}$ $*([\![\delta]\!], w) = *([\![\delta]\!], w')$, then for any T_i survivor world w_0, there is a world w_1 such that w_1 is a T survivor, provided T has a survivor. And further $* \upharpoonright (T_i, w_1) = * \upharpoonright (T_i, w_0)$, $Th(w_1 \upharpoonright Formula(T_i)) = Th(w_0 \upharpoonright Formula(T_j))$.

Proof
Suppose again w is a T_i survivor world and suppose that T_1 has a survivor world w_1. If every T_i formula is T-independent of every T_1 formula, then by the irrelevance lemma we have a survivor for T_i and T_1. Now suppose that T_i and T_1 share a common set of formulas Δ. By hypothesis the propositions p defined by the common formulas in $Ant(T_i)$ and $Ant(T_1)$ are such that any T_i survivor world w_i, $*(w_i, p) = *(w_1, p)$. Again by using the irrelevance lemma on the independent formulae in $Formula(T_j) - \Delta$, we can show that there is a survivor w_2

such that $Th(w_2 \upharpoonright Formula(T_i)) = Th(w \upharpoonright Formula(T_1))$. Again by extending the argument to all the subtheories, we get the desired result. $\Box$.

It seems that we may strengthen the irrelevance still further in the case where for each of the common formulae in Δ among the theories $T_k, \ldots, T_m$ there are survivor worlds $w_k, \ldots, w_m$ that can be ordered in such a way that $*([\![\delta]\!], w_j) \subseteq *([\![\delta]\!], w_i)$ for $i \leq j$ then a T survivor world may be constructed from any T_i survivor by following the procedure already outlined and then taking for the common predicates the strongest or smallest set of normal worlds. That a world w is a T_i survivor depends solely on the structural relationship between the survivor world and the set of normal worlds it assigns to $[\![\phi]\!]$ for $\phi \in Ant(T_i)$. By the hypothesis and the construction procedure suggested, this relationship is preserved. We formalise this observation in the following lemma, but first some definitions:

Definition: Suppose $T_1, \ldots, T_n$ are sets of $>$-sentences, such that $T = T_1 \cup \ldots \cup T_n$ and either:

(a) $Formula(T_1), \ldots, Formula(T_n)$ are T-independent (if distinct), or

(b) $T_j \ldots T_m$ have a common set of formulae Δ such that
$Formula(T_1), \ldots, Formula(T_{j-1})$,
$Formula(T_j) - \Delta, \ldots, Formula(T_m) - \Delta$,
$Formula(T_{m+1}), \ldots, Formula(T_n)$
are T-independent (if distinct), and there are survivor worlds $w_j, \ldots, w_m$ of $T_j, \ldots, T_m$ that can be ordered such that for each $\delta \in \Delta$ $*(w_i, \delta) \subseteq *(w_j, \delta)$ for $i < j$.

Then we say $T_1, \ldots T_n$ has the *chain survivor property.*

Definition: Suppose $T_1, \ldots, T_m$ are theories such that $T_1 \approx\!\!\!\mid\; \phi_1$ and $\phi_1 \in T_2$, $T_2 \approx\!\!\!\mid\; \phi_2$ and $\phi_2 \in T_3, \ldots, T_{m-1} \approx\!\!\!\mid\; \phi_{m-1}$ and $\phi_{m-1} \in T_m$ and that $\cup T_i \setminus \{\phi_1, \ldots \phi_m\} = T$. Then we say that $T_1, \ldots, T_m$ are *default expanded subtheories of* T.

Starkly Strengthened Irrelevance Lemma (SSI)
Suppose T is a theory and $T_1, \ldots, T_n$ are default expanded subtheories of T. If $Formula(T_i) - \Delta$, $Formula(T_j) - \Delta$, are T-independent for $i \neq j$ for some set $\Delta \subseteq L_{T_k} \cap \ldots \cap L_{T_m}$ and there are survivor worlds $w_k, \ldots, w_m$ for $T_k, \ldots, T_m$ respectively such that for each $\delta \in \Delta$ such that $T_k, \ldots, T_m$ have the chain survivor property with respect to δ, then for any T_i survivor world w_0 there is a world w_1 such that w_1 is a T survivor and $*(w_1, \psi) = *(w_0, \psi)$ for all independent formulae in $Formula(T_i)$, and for any common formula $\delta \in \Delta$, $*(w_1, [\![\delta]\!]) \subseteq *(w_0, [\![\delta]\!])$ and if $w_0 \in S_i \cap *(w, [\![\delta]\!])$ then $w_1 \in S_i \cap *(w, [\![\delta]\!])$.

Corollary:
Given a theory T and subtheories $T_1, \ldots T_n$ such that $Formula(T_i) - \Delta$ and $Formula(T_j) - \Delta$ are T-independent for $i \neq j$ for some set $\Delta \subseteq T_k \cap \ldots \cap T_m$ and these subtheories have the chain survivor property with respect to any $\delta \in \Delta$, then if $T_i \approx\!\!\!\mid\; \phi$, then $T \approx\!\!\!\mid\; \phi$.

The proof of this follows from the fact that for any T_i subtheory the nonmonotonic consequences are verified by the construction of a survivor world (see propositions 1 – 6 in Asher and Morreau 1991). By the starkly strengthened irrelevance lemma (SSI) we can construct a survivor world that also survives through all of the normalisation chains defined by $Ant(T)$.

We now turn to the patterns that are commonly used in this paper to explain the examples. There are three: the Cascaded Penguin Principle, the Nixon Polygon, and the Double Penguin Principle. We give a short analysis of each one within this framework.

We will make use of the following fact: If A is an admissable model of L_{DRP} in which there is a T-survivor world w then the theory of w is an element of ☺. We will plan to construct small models to verify the defeasible inference patterns that we use.

The Cascaded Penguin Principle

Consider text (2).

(2) Max fell. John pushed him.

The theory T relevant to this example consists of the following:

The Basic Axioms
$\langle \tau, \alpha, \beta \rangle > Narration(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge Cause(\beta, \alpha) > Explanation(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta) > Cause(\beta, \alpha)$

The Axioms of Temporality
$\Box(Cause(\beta, \alpha) \rightarrow \neg Narration(\alpha, \beta))$
$\Box(Narration(\alpha, \beta) \rightarrow \neg Explanation(\alpha, \beta))$

The information state S_0 that is to be normalised contains all this information. In addition it verifies: $\langle \tau, \alpha, \beta \rangle$ and $\mathit{Info}(\alpha, \beta)$. And by axiom (A4) we also have:

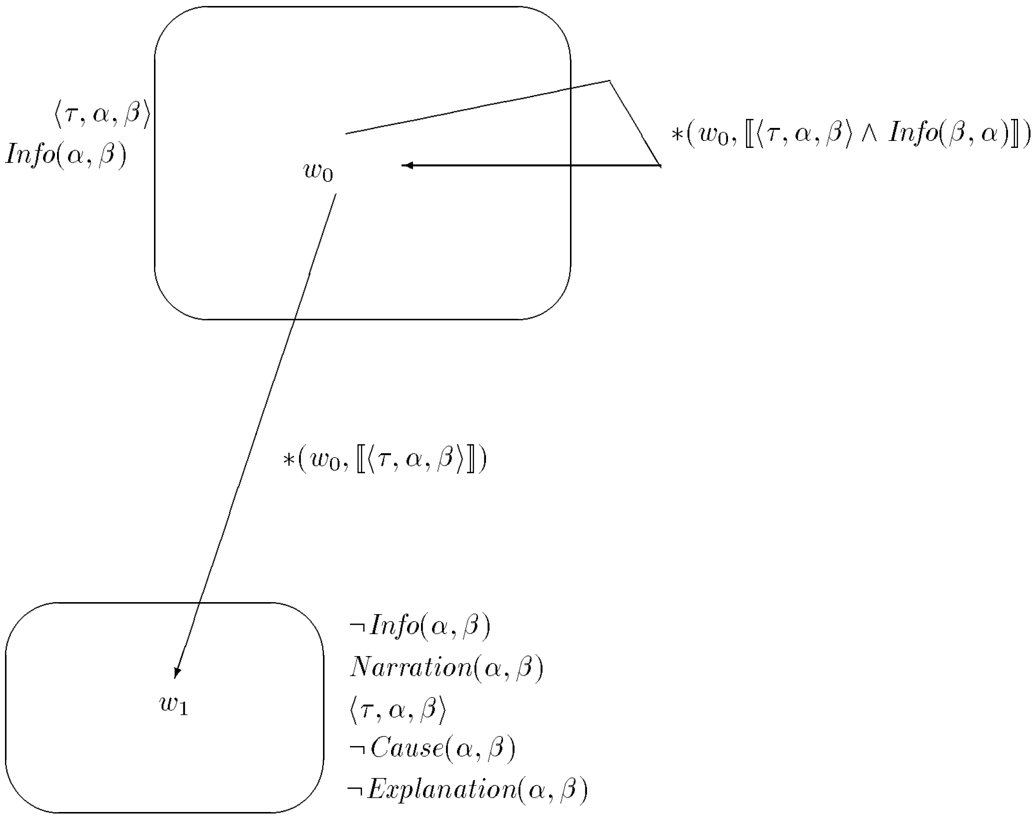
$\langle \tau, \alpha, \beta \rangle > \neg Cause(\beta, \alpha)$
$\langle \tau, \alpha, \beta \rangle > \neg \mathit{Info}(\beta, \alpha)$

We will isolate the following default expanded subtheories of T that allow us to use the SSI.

T_1:
$\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle > Narration(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta) > Cause(\beta, \alpha)$
$\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta) > \langle \tau, \alpha, \beta \rangle$
$\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta) > \neg Narration(\alpha, \beta)$
$\Box(Cause(\beta, \alpha) \rightarrow \neg Narration(\alpha, \beta))$
$\Box(Narration(\alpha, \beta) \rightarrow \neg Explanation(\alpha, \beta))$

T_2:
$\langle \tau, \alpha, \beta \rangle \wedge Cause(\beta, \alpha)$
$\langle \tau, \alpha, \beta \rangle > Narration(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge Cause(\beta, \alpha) > Explanation(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge Cause(\beta, \alpha) > \langle \tau, \alpha, \beta \rangle$
$\Box(Narration(\alpha, \beta) \rightarrow \neg Explanation(\alpha, \beta))$

T_1 and T_2 have common formulas except for $\mathit{Info}(\alpha, \beta)$, which is independent of all the others. We want to check that T_1 and T_2 have the chain survivor property with respect to their common vocabulary. Of course, all this can only be checked given the construction of an appropriate T_1 and T_2 survivor world and checking the properties. First, we turn to a T_1 survivor world. w_0 below is an appropriate T_1 survivor world.



The normal worlds above (drawn as the quadrangles) verify the formulae to their sides. It can be checked that the model above is an admissable L_{DRP} model subject to the following proviso: for any proposition $p \neq [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!]$ and $p \neq [\![\langle \tau, \alpha, \beta \rangle]\!]$, $*(w_1, p) = *(w_0, p) = p$.

We now show that w_0 is a survivor world for T_1. Typically in CE we do so by induction on the length of normalisation chains. Given that $\mathit{Ant}(T_0)$ is finite, however, the chains are finite; in fact there are just two sorts of normalisations one with $[\![\langle \tau, \alpha, \beta \rangle]\!]$ and one with $[\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!]$. The information state, S_0 verifies the axioms and the information $\langle \tau, \alpha, \beta \rangle$ and $\mathit{Info}(\alpha, \beta)$. So clearly, $w_0 \in S_0$.

Lemma 1
If $w_0 \in S$, then $w_0 \in N * (S, p)$, where $p \in \{[\![\langle \tau, \alpha, \beta \rangle]\!], [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!]\}$.

Proof
Suppose $w_0 \in S_0$ and $\phi = [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!]$. $w_0 \in *(w_0, [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!])$ and so $w_0 \in *(S, [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!])$ and so

$N * (S, [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!]) =$
$\{w \in S : w \notin [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!] \setminus *(w_0, [\![\langle \tau, \alpha, \beta \rangle \wedge \mathit{Info}(\alpha, \beta)]\!])\}$

Since $w_0 \in *(w_0, [\![\langle\tau,\alpha,\beta\rangle \wedge \mathit{Info}(\alpha,\beta)]\!])$ and $w_0 \in S$ by hypothesis, $w_0 \in N * (S, [\![\langle\tau,\alpha,\beta\rangle \wedge \mathit{Info}(\alpha,\beta)]\!])$.

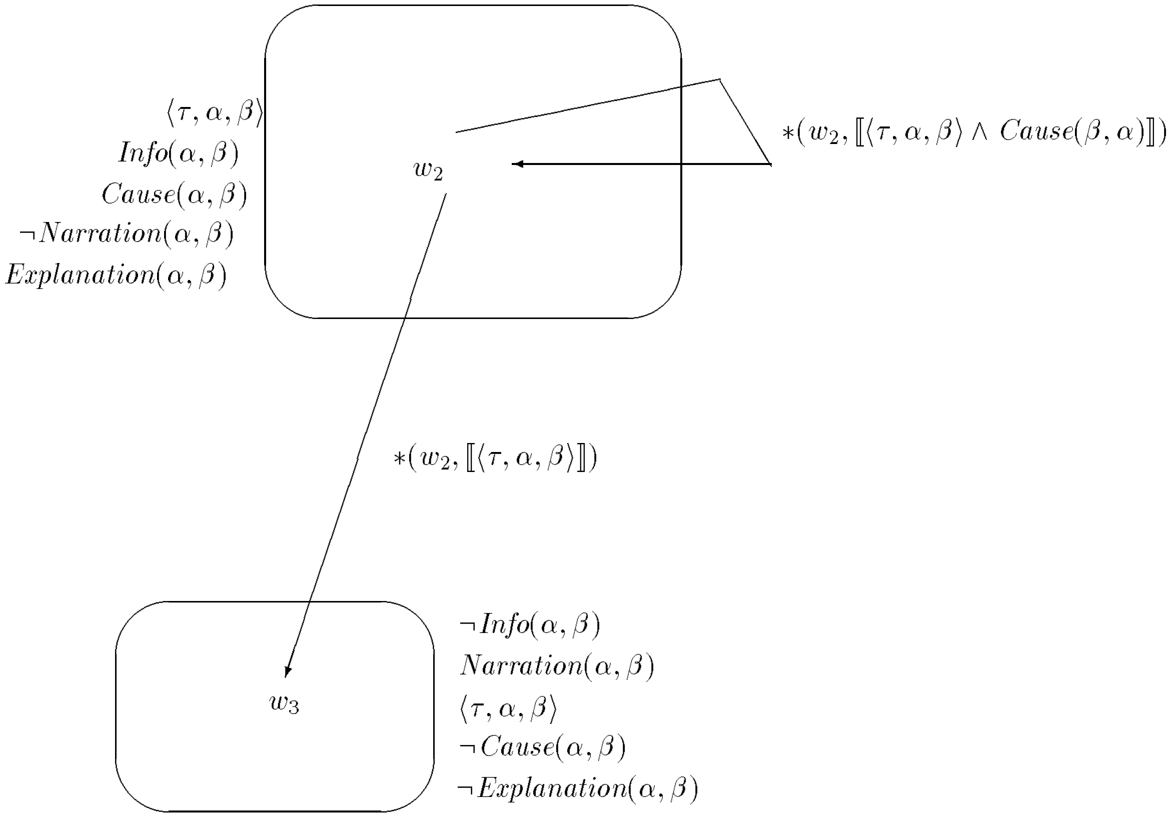
Now suppose $\phi = [\![\langle\tau,\alpha,\beta\rangle]\!]$. Then $w_0 \notin *(w_0, [\![\langle\tau,\alpha,\beta\rangle]\!])$. But $*(S, [\![\langle\tau,\alpha,\beta\rangle]\!]) \models \neg\mathit{Info}(\alpha,\beta)$ and $S \models \mathit{Info}(\alpha,\beta)$. So $S \cap *(S, [\![\langle\tau,\alpha,\beta\rangle]\!]) = \emptyset$. So $N * (S, [\![\langle\tau,\alpha,\beta\rangle]\!]) = S$ and so $w_0 \in N * (S, [\![\langle\tau,\alpha,\beta\rangle]\!])$. $\Box$

Given that w_0 is a T_1 survivor world, we now show that the fixpoint of every normalisation chain leading from S_0 verifies $\neg Narration(\alpha,\beta)$ and $Cause(\beta,\alpha)$.

Fact 2
$T_1 \mid\approx \neg Narration(\alpha,\beta) \wedge Cause(\beta,\alpha)$

Proof
Suppose that there is a chain C such that the fixpoint C_λ contains a world v such that $v \models \neg(\neg Narration(\alpha,\beta) \wedge Cause(\beta,\alpha))$. So $v \models Narration(\alpha,\beta) \vee \neg Cause(\beta,\alpha)$. $N * (C_\lambda, [\![\langle\tau,\alpha,\beta\rangle \wedge \mathit{Info}(\alpha,\beta)]\!]) = C_\lambda$. But then $v \in *(C_\lambda, [\![\langle\tau,\alpha,\beta\rangle \wedge \mathit{Info}(\alpha,\beta)]\!])$. So then $v \models \neg Narration(\alpha,\beta) \wedge Cause(\beta,\alpha)$, which contradicts our hypothesis. $\Box$

Now we turn to the appropriate T_2 survivor world w_2. Given T_1, T_2 is a default expanded subtheory of T. Here now is the normal worlds set-up for T_2.



The normal worlds for the other propositions can be filled in; only the above are relevant to making our point. This is also a DRP admissable model subject to the following proviso: for any proposition $p \neq [\![\langle\tau,\alpha,\beta\rangle \wedge Cause(\beta,\alpha)]\!]$, $p \neq [\![\langle\tau,\alpha,\beta\rangle]\!]$, $*(w_2,p) = *(w_3,p) = p$.

Suppose $S_1 \models T_2$. The proof that w_2 is a T_2 survivor world and that the fixpoints of every normalisation chain defined on S_1 goes exactly as before. But note that by the construction the common predicates for T_1 and T_2 have the chain survivor property with respect to these formulas. By SSI, there is a common survivor world for T and the fixpoint of every normalisation chain defined with respect to T verifies the conjunction of all formulas verified by each local fixpoint: I.e. $\neg Narration(\alpha, \beta)$, $Explanation(\alpha, \beta)$ and $Cause(\beta, \alpha)$.

This verifies the desired conclusions of the Cascaded Penguin.

The Nixon Polygon

Consider text (7).

(7) a. Guy experienced a lovely evening last night.
b. He had a great meal.
c. He ate salmon.
d. He devoured lots of cheese.
e. He won a dancing competition.

Let us assume a constituent α about a dinner, is elaborated by constituents β_1 and β_2 — about various courses. We now consider how to attach a constituent β_3 about winning a dance competition to this text structure. Both β_2 and α are open for β_3 and we have to look at both possibilities.

We suppose a basic information state verifying the following subtheories:

T_1:
$Elaboration(\alpha, \beta_2) \wedge \langle \tau, \alpha, \beta_3 \rangle \wedge \langle \tau, \beta_2, \beta_3 \rangle$
$\langle \tau, \alpha, \beta_3 \rangle > \neg Elaboration(\beta_3, \alpha)$
$\Box(Narration(\alpha, \beta_3) \rightarrow \neg Elaboration(\alpha, \beta_3)$

T_2:
$\langle \tau, \alpha, \beta_3 \rangle \wedge \neg Elaboration(\alpha, \beta_3) \wedge \langle \tau, \beta_2, \beta_3 \rangle \wedge Elaboration(\alpha, \beta_2)$
$\langle \tau, \beta_2, \beta_3 \rangle > Narration(\beta_2, \beta_3)$
$Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3) > \neg Narration(\beta_2, \beta_3)$
$\Box(Narration(\alpha, \beta_3) \rightarrow \neg Elaboration(\alpha, \beta_3))$

A survivor world for T_1 is not difficult to construct. But a survivor world for T_2 is more involved. Let w_0 verify $Narration(\beta_2, \beta_3)$, $\langle \tau, \beta_2, \beta_3 \rangle$ and $\langle \tau, \alpha, \beta_3 \rangle$. And let w_1 verify $\neg Narration(\beta_2, \beta_3)$, $\neg Elaboration(\alpha, \beta_3)$, $Elaboration(\alpha, \beta_2)$ and $\langle \tau, \alpha, \beta_3 \rangle$. Then we show:

Lemma 5
w_0 survives to the fixpoint of any T_2 normalisation chain in which no normalisation of the form $N * (S_n, [\![\neg Elaboration(\alpha, \beta_3)]\!])$ occurs in C prior to a normalisation of the form $N * (S_m, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!])$. And w_1 survives to the fixpoint of any T_2 normalisation chain in which no normalisation of the form $N * (S_m, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!])$ occurs before a normalisation of the form $N * (S_n, [\![\neg Elaboration(\alpha, \beta_3)]\!])$.

The proof of the two cases is completely symmetric. So let us illustrate by taking the case of w_0 first. Suppose no normalisation of the form $N * (S_n, [\![\neg \mathit{Elaboration}(\alpha, \beta_3)]\!])$ occurs in C prior to a normalisation of the form $N * (S_m, [\![\langle \beta_2, \beta_3 \rangle]\!]))$. There are only three sorts of links in a T_2 normalisation chain: $N * (S_m, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!])$, $N * (S_n, [\![\neg \mathit{Elaboration}(\alpha, \beta_3)]\!])$ and $N * (S_n, [\![Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3)]\!])$. By the specification of w_0 above $w_0 \models T_2$. If $S =$☺$+T_2$, then $w_0 \in S$. $w_0 \in *(w_0, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!])$, so $*(S, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!]) \cap S \neq \emptyset$ and so

$N * (S, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!]) = \{w \in S : w \notin [\![\langle \tau, \beta_2, \beta_3 \rangle]\!] \setminus *(S, \langle \tau, \beta_2, \beta_3 \rangle)\}$

and

$w_0 \in N * (S_m, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!])$.

Let $N * (S_m, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!]) = S'$. Suppose now that $N * (S', [\![\neg \mathit{Elaboration}(\alpha, \beta_3)]\!])$ is the next normalisation in the chain prior to the fixpoint of $N(S'', [\![Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3)]\!])$. Since

$S'' \models Narration(\beta_2, \beta_3)$ but

$S'' \models Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3) > \neg Narration(\beta_2, \beta_3)$,

$S'' \cap *(S'', [\![Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3)]\!]) = \emptyset$.

So $N * (S'', [\![Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3)]\!]) = S''$

But then $w_0 \in N * (S'', [\![Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3)]\!])$. A similar arugment shows that w_0 survives if the normalisation chain with respect to:

$[\![Elaboration(\alpha, \beta_2) \wedge \neg Elaboration(\alpha, \beta_3)]\!]$

comes prior to the normalisation chain with respect to $[\![\neg \mathit{Elaboration}(\alpha, \beta_3)]\!]$. Since this exhausts all the types of T_2 normalisation chains in which no normalisation of the form $N * (S_n, [\![\neg \mathit{Elaboration}(\alpha, \beta_3)]\!])$ occurs in C prior to a normalisation of the form $N * (S_m, [\![\langle \tau, \beta_2, \beta_3 \rangle]\!])$, we have the desired result: w_0 survives in every T_2 normalisation chain in which no normalisation of the form $N * (S_n, [\![\neg \mathit{Elaboration}(\alpha, \beta_3)]\!])$ occurs in C prior to a normalisation of the form $N * (S_m, [\![\langle \beta_2, \beta_3 \rangle]\!])$. $\square$

By Lemma 5, we see that in some fixpoints w_0 survives: in others w_1. By an argument parallel to that in Fact 2, we conclude then that at those fixpoints C_λ in which w_0 survives $C_\lambda \models Narration(\beta_2, \beta_3)$ and at those fixpoints C'_λ in which w_1 survives $C'_\lambda \models \neg Narration(\beta_2, \beta_3)$. So we conclude:

Fact 6
$T_2 \not\approx Narration(\beta_2, \beta_3))$

But further, since no other axiom of T_2 allows us to conclude any other discourse relation between β_2 and β_3, we have that $S =$☺$+T_2$ is such that $S \models \langle \beta_2, \beta_3 \rangle$ and $S \not\models R(\beta_2, \beta_3)$ for any discourse relation R. We assume such a state S to be incoherent.

Because we must avoid whenever possible incoherent information states, we seek another attachment. We thus look at the information state S''' that is the update of ☺ with T_3:

T_3:
$\langle \tau, \alpha, \beta_3 \rangle$
$\langle \tau, \alpha, \beta_3 \rangle > Narration(\alpha, \beta_3)$

An easy survivor argument shows that:

$T_3 \approx\models Narration(\alpha, \beta_3)$

Double Penguin Principle

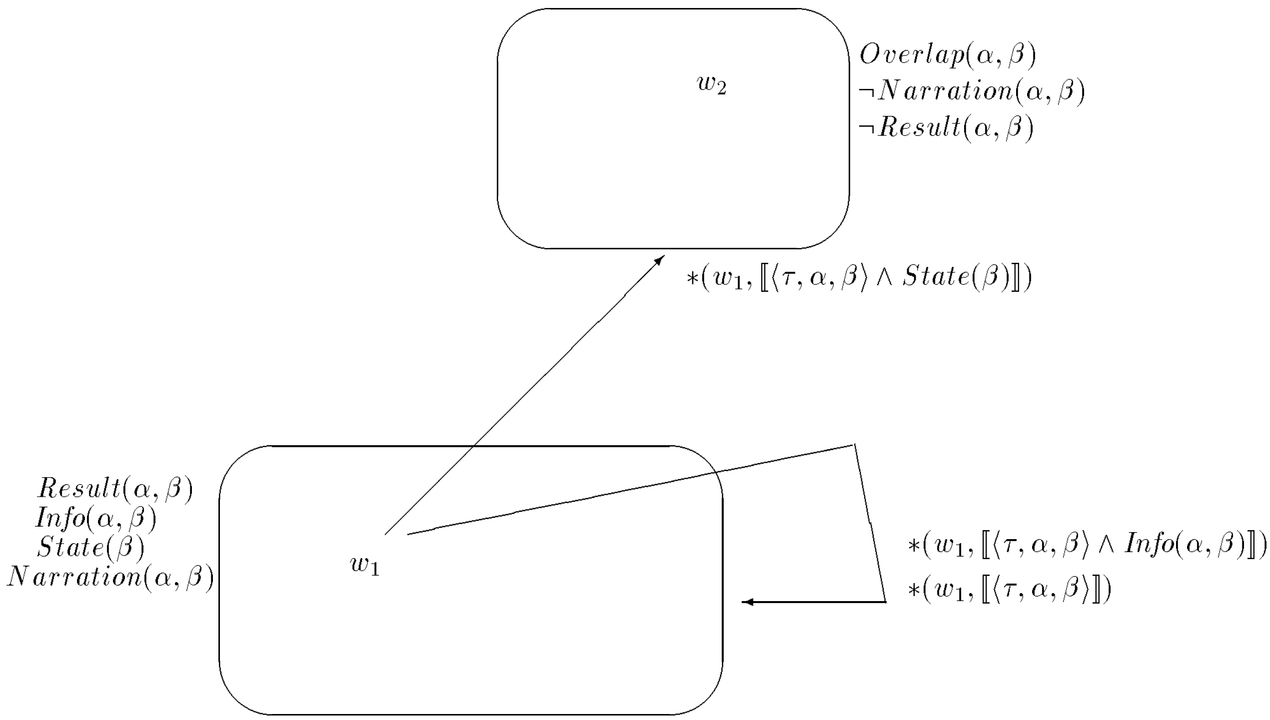
The relevant example for the Double Penguin is text (4).

(4) Max switched off the light. The room was dark.

The axioms for this example involve two resolvable conflicts — one between States Overlap and Narration and one between States Overlap and Result.

T:
$\langle \tau, \alpha, \beta \rangle > Narration(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge Info(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge State(\beta) > Overlap(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge State(\beta) > \langle \tau, \alpha, \beta \rangle$
$\langle \tau, \alpha, \beta \rangle \wedge Info(\alpha, \beta) > \langle \tau, \alpha, \beta \rangle$
$\langle \tau, \alpha, \beta \rangle \wedge Info(\alpha, \beta) > Cause(\alpha, \beta)$
$\langle \tau, \alpha, \beta \rangle \wedge Cause(\alpha, \beta) > Result(\alpha, \beta)$
$\Box(Overlap(\alpha, \beta) \rightarrow \neg Narration(\alpha, \beta))$
$\Box(Overlap(\alpha, \beta) \rightarrow \neg Reuslt(\alpha, \beta))$
$\Box(Info(\alpha, \beta) \rightarrow state(\beta))$

Because of the connections between the various formulas in $Ant(T)$ — e.g. between $\langle \tau, \alpha, \beta \rangle$, $\langle \tau, \alpha, \beta \rangle \wedge Info(\alpha, \beta)$ and $\langle \tau, \alpha, \beta \rangle \wedge State(\beta)$, subtheories that each do one of the Penguin inferences will not have the required chain survivor property with respect to their common vocabulary. So we require one large survivor proof. This proof depends on the following T survivor:



Again, analogous proofs to those in Lemmas 2 – 5, we may show:

Lemma 8
w_1 is a T survivor

and from this follows:

Fact
$T \mathrel{|\!\approx} Result(\alpha, \beta) \wedge Narration(\alpha, \beta) \wedge \neg Overlap(\alpha, \beta)$

References

Aqvist, L [1972] Logic of Conditionals, *Journal of Philosophical Logic.*

Asher, N. [in press] *Reference to Abstract Objects in English: A Philosophical Semantics for Natural Language Metaphysics*, Kluwer Academic Publishers.

Asher, N. and Morreau, M. [1991] Common Sense Entailment: A Modal Theory of Non-monotonic Reasoning, in *Proceedings to the 12th International Joint Conference on Artificial Intelligence*, Sydney Australia, August 1991.

Caenepeel, M. [1991] Event Structure vs. Discourse Structure, in *Proceedings of the* DANDI *Workshop on Discourse Coherence*, Edinburgh April 1991.

Chellas, B. F. [1980] *Modal Logic.* Cambridge: Cambridge University Press.

Cohen, P. R., Morgan, J. and Pollack, M. E. [1991] (eds.) *Intentions in Communication*, MIT Press, Cambridge Massachusetts USA.

Dahlgren, K. [1988] *Naive Semantics for Natural Language Understanding*, Kluwer Academic Publishers, Dordrecht, Holland.

Dowty, D. [1986] The effects of aspectual class on the temporal structure of discourse: semantics or pragmatics? *Linguistics and Philosophy*, **9**, 37-61.

Grice, H. P. [1975] Logic and Conversation. In Cole, P. and Morgan, J. L. (eds.) *Syntax and Semantics*, Volume 3: *Speech Acts*, pp41-58. New York: Academic Press.

Grosz, B. J. and Sidner, C. L. [1986] Attention, Intentions, and the Structure of Discourse. *Computational Linguistics*, **12**, 175-204.

Hinrichs, E. [1986] Temporal anaphora in discourses of English. *Linguistics and Philosophy*, **9**, 63-82.

Hobbs, J. R. [1979] Coherence and Coreference. *Cognitive Science*, **3**, 67-90.

Hobbs, J. R. [1985] On the Coherence and Structure of Discourse. Report No. CSLI-85-37, Center for the Study of Language and Information, October, 1985.

Joshi, A. Webber, B. and Weischedel, R. [1984] Default reasoning in interaction. In *Proceedings of the Non-Monotonic Reasoning Workshop*, AAAI, New York, October, 1984, pp144-150.

Kamp, H. [1991] The Perfect and Other Tenses in French and English, in Kamp, H. (ed.), *Tense and Aspect In English and French*, DYANA deliverable 2.3B, available from the Centre for Cognitive Science, University of Edinburgh.

Kamp, H. and Reyle, U. [in press] *From Discourse to Logic; Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Kluwer Academic Publishers, Dordrecht Holland.

Kamp, H. and Rohrer, C. [1983] Tense in texts. In Bauerle, R., Schwarze, C. and von Stechow, A. (eds.) *Meaning, Use and Interpretation of Language*, pp250-269. Berlin: de Gruyter.

Konolige, K. [1988] Hierarchic Autoepistemic Theories for Nonmonotonic Reasoning: Preliminary Report, Technical Note No. 446, SRI International, Menlo Park, August 1988.

Lascarides, A. [1988] The Semantics of Aspectual Classes Using Reference to Intervals, Research Report, EUCCS/RP-22, Centre for Cognitive Science, University of Edinburgh.

Lascarides, A. [1992] Knowledge, Causality and Temporal Representation. *Linguistics*, volume **30**, number 5.

Lascarides, A. and Asher, N. [in preparation] The Pluperfect in Narrative Discourse, research report.

Lascarides, A. Asher, N. and Oberlander, J. [1992] Inferring Discourse Relations in Context, in *Proceedings of the 30th Annual Meeting of the Association for Computational Linguistics*, Delaware June 1992.

Lascarides, A. and Oberlander, J. [1992] Temporal Coherence and Defeasible Knowledge, *Theoretical Linguistics*, **18**.

Lifschitz, V. [1984] Some results on circumscription, in *Proceedings of the Non-Monotonic Reasoning Workshop*, AAAI, New York, October 1988, pp. 151–164.

McCarthy, J. [1980] Circumscription: A Form of Nonmonotonic Reasoning, *Artificial Intelligence*, **13**, pp 27–39.

Moens, M. and Steedman, M. J. [1988] Temporal ontology and temporal reference. *Computational Linguistics*, **14**, 15-28.

Moore, R. C. [1984] A Formal Theory of Knowledge and Action, Technical Note, SRI International Number 320, February 1984.

Morreau, M. [1992] Norms or Inference Tickets? A Frontal Collision Between Intuitions, in Briscoe, E. J., Copestake, A., and de Paiva, V. (eds.) *Default Inheritance in Unification-based Approaches to the Lexicon*, Cambridge University Press.

Partee, B. [1973] Some structural analogies between tenses and pronouns in English, *Linguistics & Philosophy*, Volume **70** number 18, pp. 601–609.

Partee, B. [1984] Nominal and Temporal Anaphora *Lingiustics & Philosophy*, Volume **7** pp.243–286.

Polanyi, L. [1985] A Theory of Discourse Structure and Discourse Coherence. In Eilfort, W. H., Kroeber, P. D. and Peterson, K. L. (eds.) *Papers from the General Session at the Twenty-First Regional Meeting of the Chicago Linguistics Society*, Chicago, April 25-27, 1985.

Reichenbach, H. [1947] *Elements of Symbolic Logic*. London: Macmillan.

Reiter, R. [1980] A Logic for Default Reasoning, *Artificial Intelligence*, **13**, pp81–132.

Scha, R. and Polanyi, L. [1988] An augmented context free grammar. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, pp573–577.

Thompson, S. and Mann, W. [1987] Rhetorical Structure Theory: A Framework for the Analysis of Texts. In *IPRA Papers in Pragmatics*, **1**, pp79–105.

Veltman, F. [1990] Defaults in Update Semantics, DYANA deliverable 2.5a, available from Centre for Cognitive Science, University of Edinburgh.

Webber, B. [1991] Structure and Ostension in the Interpretation of Discourse Deixis, *Language and Cognitive Processes*, Volume **6** Number 2, pp 1-7–135.