

Analysing Aviation Accidents Using WB-Analysis – an Application of Multimodal Reasoning

Peter Ladkin, Karsten Loer
Technische Fakultät, Universität Bielefeld
33501 Bielefeld, Germany
{ladkin | karlo}@rvs.uni-bielefeld.de
<http://www.rvs.uni-bielefeld.de>

Abstract

We describe our ongoing work in accident analysis. Accident reports should tell us at least what the accident was and what the critical events were. A third requirement they should fulfil is to *explain* these events (see below) and their sequence (\rightarrow temporal reasoning). Explanation concerns causes (\rightarrow causal reasoning), human intentions, purposes, capabilities and behavior (so-called *human factors*). Causality also involves the unfolding of events in time (Ladkin, *Explaining Failure in Tense Logic*, RVS-RR-96-13). We include social factors - obligations and the regulatory environment - amongst the human factors (\rightarrow deontic reasoning). Our goal is a rigorous method of incident explanation which contains search procedures for relevant facts and insists on rigorously formal proofs of an explanation's correctness and relative sufficiency.

Reasoning about accidents: the basics

We describe *Why-Because Analysis* (WBA), a method of deriving explanations of incidents and accidents, and of rigorously proving the resulting explanations correct according to certain formal criteria. To our knowledge, this is the first system which accomplishes both these goals. The logic for the formal proofs is called *Explanatory Logic* (EL), and we perform correctness proofs in EL by hand in the hierarchical style advocated by Lamport. WBA is explained in full in the monograph [Ladkin, Loer 1998].

1. We must determine what to reason about (what's in the universe?). To do this in a formally correct way, we describe the world in the ontology of TLA, Lamport's Temporal Logic of Actions, because this ontology has shown itself sufficient for describing the temporal behavior of artifacts (process algebra & Petri net semantics use a similar ontology).
2. In principle, this should involve us in clarifying the relation between set theory (or other formal data structuring) and 'the world'. But we leave this to philosophers of mathematics. We're more interested in the accidents.

3. We must determine what kind of reasoning is involved, and encode it in inference rules in a formal logic sufficient for proving correctness of the analysis.

Since a narrative is involved, tense logic (with the Kripke semantics) is an appropriate reasoning tool, given that the ontology has been declared suitable. The ontology of TLA (Lamport's Temporal Logic of Actions) is sufficient for (A) description of machine behavior, (B) formulation of accident histories, and (C) determination of sequences of states leading to an accident. *States* are individuated by the collection of state predicates which are true in that state. This makes states into *types*. Particular occurrences of states can be identified by means such as timestamps or positions in the causal chain. States furthermore may have a duration. In contrast, *events* are particulars, representing specific changes in state. An individual event cannot recur, but its type (a TLA *action*) can be instantiated more than once. There are also *processes*, state/event mixes of bounded duration which describe undifferentiated actions. *Non-events*, the non-occurrence of awaited events, are also important (see below).

The logical operators in TLA are insufficient by themselves for adequate reasoning about accidents. A causal relation $\square\rightarrow$ (more exactly, a relation of *causal explanation*) is required. We have argued elsewhere that causal relations cannot be defined in pure tense logic (Ladkin, *Some Dubious Theses...*, RVS-RR-96-14), a thesis accepted by philosophers but apparently not by some computer scientists.

How then to handle causality?

The most appropriate semantics for our purpose is the Lewis criterion for causality (Lewis 1973a), based on Lewis's formal semantics for counterfactual conditionals (Lewis 1973b):- Let A and B be *events* or *state instances*. Then, informally,

$$A \text{ is a causal factor of } B \equiv \left(\begin{array}{l} A \text{ and } B \text{ both occurred, and} \\ \text{in the nearest possible worlds} \\ \text{in which } A \text{ did not happen,} \\ \text{neither did } B. \end{array} \right)$$

Causal factors usually succeed each other in a tempo-

ral order. (This temporal order dominates the common narratives of accidents as noted in [Ladkin, Loer 1998]). Using J.S. Mill's criterion for causality, the so-called *Method of Difference*, we find that temporal succession is (at least) a hint towards causality. This leads directly to the axiom:

Axiom 1 $\vdash (A \Rightarrow^* B) \Rightarrow (A \leftrightarrow B)$

An n-fold succession of causal-factor relations between chained factors implies a temporal succession between (at least) the first and the last factor of this chain.

in which the relation \leftrightarrow is that of temporal succession and the relation \Rightarrow^* is intended to denote the transitive closure of \Rightarrow , 'causal factor of', the primary causal operator advocated by Lewis in (1973a) and defined by him in terms of the counterfactual conditional from (Lewis 1973b). Although one cannot define the transitive closure of a relation from that relation in first-order logic, one can nevertheless axiomatise it effectively in the standard manner used by logicians, which we do. The relation \Rightarrow^* is, according to Lewis, true causality. Nothing we do hangs on this particular identification, however.

Using *modus ponens*, Axiom ?? leads to a derived inference rule we will need for our analysis:

$$\frac{A \Rightarrow^* B}{A \leftrightarrow B} \quad (1)$$

This rule means that causalities must be consistent with temporal order.

Our investigation method, *Why-Because Analysis*, seeks to reverse this order, like proofsearch. The Lewis criterion for being a causal factor is formulated as:

$$A \Rightarrow B \triangleq (A \Box \rightarrow B) \wedge (\neg A \Box \rightarrow \neg B) \quad (2)$$

which leads to the two-way inference rule:

$$\frac{\frac{A \Box \rightarrow B}{\neg A \Box \rightarrow \neg B}}{A \Rightarrow B} \quad (3)$$

The semantics of $\Box \rightarrow$ is a possible-world semantics (Lewis 1973b). This '*nearest possible world*' relation is Lewis's extension to Kripke semantics. He introduces an additional relation of nearness:

World X is at least as near as world Y to world W . Even though Lewis has given complete sets of inference rules for $\Box \rightarrow$, during WBA we find ourselves mostly evaluating the truth of assertions involving $\Box \rightarrow$ by using semantic arguments within the "nearest possible world" semantics. Formally, the semantics is used as follows. Fix W for the moment. The nearness relation yields a binary relation \preceq_W , namely

$$X \preceq_W Y \quad (4)$$

The relation \preceq_W is axiomatised by Lewis as a *total preorder*. That means that any two worlds can be compared in terms of their similarity to world W ; either the one or the other is more similar, or they are both equally similar. The Lewis semantics for $\Box \rightarrow$ is that

$A \Box \rightarrow B$ in a world W if and only if B is true in all the nearest worlds to W in which A is true.

The case we use is $W = \text{the actual world}$.

One defines the concept 'nearest': a world X is nearest to a world W if and only if, for all worlds Y , world X is at least as near as world Y to world W . Suppose A is true in world W . Then the set of nearest worlds to W in which A is true consists of precisely W itself. Then $A \Box \rightarrow B$ is true in W just in case B is also true. Thus we have the rule:

$$\frac{\frac{A}{B}}{A \Box \rightarrow B} \quad (5)$$

which allows us to reduce the Lewis criterion for counterfactuals in the form in which we use it, to explain the causal-factor relation between facts A and B rather than fictions, to

$$\frac{\frac{A \wedge B}{\neg A \Box \rightarrow \neg B}}{A \Rightarrow B} \quad (6)$$

Other significant points are:

1. Lewis semantics is more easily seen as a relation of *causal explanation* rather than *causality*, since it includes purely logical notions (observe that A is a *causal factor of B* if A is a *logical consequence of B*); and an event may indeed have a singular cause, even though explaining why that event happened may require the invocation of many factors (Davidson 1967, Lewis 1973a). Hence our choice of name as Why-Because Analysis.
2. Lewis's semantics can have a technical problem with handling *overdetermination*: the occurrence of two or more independent causal factors B and C for A that are individually sufficient. We finesse this problem where necessary by identifying and handling such cases individually;
3. we have found the Lewis semantics alone adequate for detecting common forms of reasoning error, for example:
 - (a) Cali: communication between crew and ATC as well as the ID/freq identity of ROZO and ROMEO shown to be causal factors
 - (b) Warsaw: placement of the ridge and the runway surface shown to be causal factors
 - (c) O'Hare: failure of the slat-retraction indicating system and the inability of the crew to observe the wing from the cockpit shown to be causal factors

All these points, although occurring in the respective accident reports, were not cited as causal factors therein.
4. We have performed full informal WB-analyses of the Cali, Warsaw and Nagoya accident reports, and a full formal analysis with proof of correctness of an incident in which an aircraft landed at the wrong airport.

5. Since the Lewis relation is binary, we can represent results of the analysis in a graph whose nodes are events/states/processes, called the *Why..Because..-Graph* (WB-Graph). This graph has a textual form (e.g., Figure ??) as well as its graphical form (Figure ??). We find both useful.
6. We can determine the 'presence' of causally-explanatory non-events in a WB-analysis by comparing the actual pattern of (occurring) events against rule-based 'standard operating procedures' (SOPs) (first suggested during joint work with Ev Palmer). This requires formalising SOPs as TLA modules [Ladkin and Loer 1998], and noting a conflict between what did happen and what *should have* happened, which brings us now to deontics.

```
[1] /* AC lands at Brussels RWY 25 */
[-.1] /* CRW opts to continue landing */
<-.2> /* AC near Brussels Airport */
<- .3> /* AC is in landing phase*/

[1.1] /\[-.1] /* Crew (CRW) realizes they are landing at the
           wrong airport */
      /\<-.2> /* CRW has safety reasons for continuing
           landing */
      /\<- .3> /* Standard Operating Procedures */

[1.1.1] /\[-.1] /* CRW gets visual contact to Brussels
           airport */
        /\{-.2> /* CRW notices that Brussels' airport
           layout is different from Frankfurt's */

[1.1.1.1] /\[-.1] /* AC breaks out under clouds, */
          /\<-.2> /* CRW procedures */
          /\<2> /* AC in BATC area */
          /\<1.2>

[...]

<1.2> /\(-.1) /* CRW did not realize that they were on wrong
           course, UNTIL: [111] */
      /\<-.2> /* AC cleared to BATC according to ATC
           procedures */

(1.2.1) /\(-.1) /* CRW addresses BATC controller as
           'Frankfurt' several times, */
        /\<-.2> /* ILS has different frequency for
           Frankfurt. */
        /\[-.3] /* CRW asks for the Bruno VOR's
           frequency. */
        /\(-.4) /* Brussels did not question the
           addressing error although it happened
           more than once */
        /\<-.5> /* BATC procedures deviate from SOPs */
        /\<1.1.1.1.2>
```

Figure 1: Excerpt from the NW Flight 052 incident analysis' (textual) WB-Graph. Notation: [X] denotes an event, <X> a state, {X} a process and (X) a non-event.

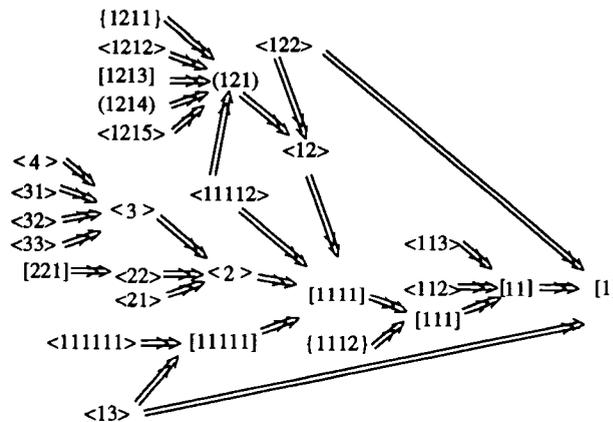


Figure 2: The pictorial version of the WB-Graph

Deontics and Alethics

Aviation is strongly determined by regulations (FAR, ICAO stans, ...) and fixed behaviors (AIM, SOPs, ...). SOPs for pilots, airlines, ATC, FAA alike can be formulated as deontics. However, this can highlight conflicting explanatory requirements!

As an example we use a recent accident, anonymised because the official investigation is not yet complete:

A large commercial transport aircraft flew into terrain at night in rainy weather while approaching the airport on a non-precision approach. Usually, a precision approach is available, but the 'glideslope' equipment had been taken out of service for upgrade. The Lewis causality criterion determines the missing glideslope equipment, night, and weather conditions to be (actual) causal factors in the accident history. However, analysed deontically, these factors play no substantial role (assuming that weather is found to have been 'normally bad') since:

1. pilots and airline knew before the flight took off that glideslope guidance was unavailable, and that they the landing would use a different approach procedure. They chose to accept this constraint (deontic). All required systems for this different procedure were available and working (as far as we know).
2. the approach was designed for conditions of reduced visibility, including night and cloudy/rainy weather of the sort supposed to have been present on the occasion of the accident
3. the weather was not optimal, but it is not known to have played a significantly unusual role (cf windshear).

Summary: The circumstances for the landing were generally known before start and with their decision to start, the CRW (and the airline) accepted them. Such deontic concerns take precedence in accident explanations over some purely physical factors such as the normal circumstances of landing.

The importance of the deontics is emphasised by (Reason 1989):

Data from the Institute of Nuclear Power Operations (INPO) shows that 92% of errors were man-made; that only 8% of the total were initiated by the operators. "The majority had their origins in either maintenance-related activities, or in fallible decisions taken within the organisational and managerial domains" (Reason 1989). These are the 'latent errors'. The Australian BASI uses Reason's latent-error model for all flight accident reports (Reason 1990).

We conclude that such deontic reasoning introduces a type of factor, human obligations arising from judgments, which takes explanatory priority over (physically) causal factors which do not fit in with the regulatory environment (formulated as obligations). Furthermore, deontic and causal reasoning interacts incompatibly - but we lack an explanation of how!

Leaving aside for now this interaction, it turns out that one can define the 'standard' deontic modal operator from $\Box \rightarrow$, by way of an alethic operator \Box_a which we need for defining the S5-type strict implication \succ . We use the deontic operator in determining the presence of factors or human failures of a sort often found in accident reports:

"X did not perform action Y."

How can one possibly determine that a non-event - the absence of something - is causally important? Formally, a *non-event* is a state. Something we await - according to our knowledge about the situation the system is in as well as the obligations following from the procedures which govern the system - does not occur. That kind of argumentation uses deontic reasoning: the procedures *ought* to be followed. Therefore the event *ought* to have happened. But it *did not*. And so we explicitly remark it. To see how we may capture this formally, consider the following principle:

We assert the existence of a non-event, given procedures *Proc* in a situation *S* if, given *S* and always *Proc*, that the procedures are continually followed, the event must necessarily occur, either then or later; *but in fact it doesn't*.

This involves two modalities, in the technical sense of the term in modal logic: *necessity* and *tense*. Because tense logics and logics of necessity are often considered separately, the same notation is used for both: we need to distinguish notation. We shall use \succ for the Lewis-Langford (Lewis, Langford 1932) S5 relation of *strict implication*: $A \succ B$ if *B* necessarily follows as a matter of logic from *A*. $A \succ B$ is definable from an alethic modality \Box_a as $\Box_a(A \Rightarrow B)$; and \Box_a is as it turns out definable from $\Box \rightarrow$. We use the plain \Box , \Diamond for the *always* and *eventually* operators of simple linear-time tense logic.

The fourth modality is obligation. The deontic axiom says that procedures *ought* to be followed:

Axiom 2 $\vdash O(Procedures)$

Suppose an event is a necessary consequence of following procedures in the given situation. Since the procedures ought to be followed, the event ought to occur: it should occur. First, how do we say *necessary*

consequence? Since we're talking about procedures or systems with behavior, we can use TLA along with the following axiom:

$$\frac{(\vdash_{TLA} A \Rightarrow B)}{A \succ B} \quad (7)$$

This can be used as a proof rule in hierarchical proofs in the following way: to prove $A \succ B$, the proof proceeds according to the proof of $A \Rightarrow B$ using the proof rules of TLA as given in, say (Ladkin, *Using the Temporal Logic of Actions - A Tutorial on TLA Verification*, RVS-RR-97-08). Given this rule for \succ , we may now formulate the *Deontic Rule* which says that when the occurrence of an event is a necessary consequence of procedures, that the event ought to happen:

$$\frac{(Hypotheses \wedge \Box Procedures) \succ \Diamond Event}{(Hypotheses \wedge O(Procedures)) \Rightarrow O(\Diamond Event)} \quad (8)$$

It follows trivially as a derived rule from Rule ?? and Axiom ?? that

$$\frac{Hypotheses}{(Hypotheses \wedge \Box Procedures) \succ \Diamond Event} \quad (9)$$

$$O(\Diamond Event)$$

The event may not in fact occur, even though it should have, because it is perfectly possible that the procedures weren't followed and thus allowed the event not to occur. In our analysis, we need to remark and reason with these events that should have occurred but didn't. We call them non-events. What kinds of objects are they? Well, non-events persist: the system state does not change in the relevant way because the event that causes this change does not occur, so non-events describe *states* whose occurrence is inferred from our knowledge of procedures, and of the current situation.

However, it is difficult to formulate this final step, the existence of non-events, as a formal inference rule, because it really tells us explicitly to remark a particular fact. We really have a meta-rule:

Axiom 3 MetaAxiom: *Explicitly add to the history those states ($\neg E$) in which E is an event, $O(E)$ is derivable, and E does not occur.*

Sufficiency of explanation:

One could search for necessary and sufficient causal explanations *A*, where *A* is a conjunction of factors, of a state or event *C*, by looking for factors which fulfil the definition:

$$A \Box \Rightarrow B \triangleq \left(\begin{array}{l} \wedge A \Rightarrow B \\ \wedge \neg B \Box \rightarrow \neg A \end{array} \right) \triangleq \left(\begin{array}{l} \wedge A \Box \rightarrow B \\ \wedge \neg A \Box \rightarrow \neg B \\ \wedge \neg B \Box \rightarrow \neg A \end{array} \right) \quad (10)$$

but it turns out this definition is too strong. Consider the operation of a FSM according to the specification

Spec. Let *Hyp* be the set of facts giving the current state of the machine. It suffices as an explanation of event or state *E* that, provided that the *Spec* is in fact followed by the machine:

$$Hyp \wedge Spec \succ E$$

This observation leads to the rule:

$$\frac{\begin{array}{l} Hyp \\ Spec \\ (Hyp \wedge \Box Spec) \succ E \\ E \end{array}}{Hyp \wedge Spec \Box \Rightarrow E} \quad (11)$$

(We distinguish *Hyp* and *Spec* because of their logical form – *Spec* and *SOPs* usually have the form $\Box A$, whereas *Hyp*, being a set of contingent and sporadic facts, will normally not).

However, the derived rule resulting from Definition ?? above would require the extra hypothesis:

$$\neg(Hyp \wedge Spec) \Box \rightarrow \neg E$$

which may not be true in particular cases in which *Spec* is far stronger than the minimal condition on the device which entails *E*. Searching for this minimal condition is often futile, often merely an interesting logical problem which is not so interesting for explaining accidents.

The same Rule ?? applies for actions affected by human operators when they have followed *SOPs* correctly, but now with the definition of *SOP* replacing *Spec*. This rule plays a substantial role in formal proofs of sufficiency of an explanation in WBA. The technical advantage of the rule with the extra hypothesis derived from (??) would be that the operator $\Box \Rightarrow$ would be definable from $\Box \rightarrow$, as are \Box_a and *O*, and thus fall under the soundness and relative completeness theorem of (Lewis 1973b). But then we wouldn't be able effectively to reason that specifications and *SOPs*, when followed, yield explanations of the occurrence of certain states and events. We must assign priority to encoding explanatory reasoning as it actually is, and admit that the application doesn't allow us quite as clean as a logic as might be wished for.

Closed World Assumptions and other Non-Monotonicity:

CWA: Accident reports use a closed-world assumption, namely that either all the significant events and states are known, or those that are not known are known to be not known. Both the CWA and other non-monotonic reasoning can be expressed in the ontology introduced above.

...and other Non-Monotonicity: In principle, the 'world' consists only of states or events obtained *directly* from instruments like cockpit voice recorder (CVR) and digital flight data recorder (DFDR, 'black box'); photographs; on-site investigation of wreckage; states, events or processes derivable by temporal, causal and deontic reasoning from these. Formally, for every 'new'

node (representing new knowledge of one of these states, events or processes) we introduce in our analysis, we have to check whether former reasoning is still valid (there are thus two cases: simple incompleteness and non-monotonicity - see below). Whenever we make an *assumption* about a cause for a state/event/process, we limit the explanatory power of the system to explanations which fulfil this assumption. To keep this limitation within bounds (we prefer to base analysis on formal argumentation rather than speculation), it would make sense formally to clone the 'existing' world before we introduce the new information, as in the method of semantic tableaux. We would need to control the potential exponential growth of the number of worlds to consider. Alternatively, we can be content with justifying 'reasonable' assumptions and ignore alternatives, but we may have to be prepared to revise these in light of further discovery (non-monotonicity). Examples:-

Call (incompleteness, monotonic reasoning):

DFDR recordings show that the machine turned left for 90 seconds. This could not be explained, until an undamaged FMC was discovered and its non-volatile memory decoded. In this case, the WB-method would yield an incomplete, but causally correct graph, which contains all information discovered, but not including grounds for the left turn. The additional information gleaned from the FMC several months after the accident can be introduced to 'complete' the graph. Such 'completions' result in additional subgraphs, but do not change the rest of the graph.

Lauda Air, Thailand (assumption, non-monotonic):

Evidence from CVR that reverse thrust (RT) was 'deployed'; but there's an interlock.

Conclusion: upset cannot be directly explained. Subsequently found a failure mode of the interlock, which in principle could allow RT to actuate in flight. Report contains no probable cause, but considers this to be a likely scenario.

Mont Ste. Odile, Strasbourg (assumption, non-monotonic):

Autopilot modes not available on DFDR; flight path shows rapid descent starting exactly at FAF. Descent rate in fpm is almost identical with required flight path angle in degrees; also the autopilot descent mode would have been engaged at FAF, where divergent behavior started. Autopilot mode control is unlabelled toggle; mode annunciation is via small letters, rate/angle larger figures. Again, this 'likely cause' is presumed.

Summary: All accident reports make a CWA: the relevant facts are those we know plus those we know we don't know. Assumptions about 'likely happenings' introduce either an extra (formal) modal dimension or non-monotonicity.

Contrastive Explanation

Contrastive explanation concerns the explanation of facts of the form *why P rather than Q occurred*, and it is a major search device in WBA for explanatory facts. Lewis (Lewis 86)[pp229-230] suggests this may be accomplished by giving information about the causal history of *P* that would not have applied to the history

of *Q*. Lipton (Lipton 91)[p42] notes that this criterion allows for unexplanatory causes. J.S. Mill's *Method of Difference* (Mill 1973)[III.VIII.2] relies on the principle that a cause must lie among the antecedent differences between a case in which the effect occurs and a case in which it does not. Mill notes that this works best with *diachronic* (before/after) contrasts. Lipton (Lipton 1991)[p43] proposes the *Difference Condition*:

To explain why P rather than Q, we must cite a causal difference between P and not-Q, consisting of a cause of P and the absence of a corresponding event in the case of not-Q.

(Lipton considers here that only events may be causes. We are considering causal factors to include nodes of all types, so relevant modifications must be made to this expression of the Difference Condition.) We apply these principles of contrastive explanation during our search in WBA for relevant facts.

In summary, we have found multimodal reasoning to be essential for formal accident analyses and their correctness proofs:

Method	Used for:
modal logic/Tense Logic	temporal reasoning
Lewis counterfactuals	causal explanation
alethic reasoning	operations according to specs and procedures
deontic reasoning	SOP violations, regulatory environment, significant non-events, 'latent' errors

References

Davidson, D., 1967. Causal Relations. *Journal of Philosophy* 64:691-703, also in Davidson, D., *Essays on Actions and Events*, Oxford U. P., 1980.

Ladkin, P. B., and Loer K., 1998 *Why-Because Analysis: The Formal Logic of Failure*. RVS-Bk-98-01, in preparation.

Lewis, D. 1973a. Causation. *Journal of Philosophy* 70:556-67. Also in Lewis, D., *Philosophical Papers Vol.II*, Oxford U. P., 1986.

Lewis, D. 1973b. *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, C.I., Langford, C.H. 1932. *Symbolic Logic*, New York: Dover Publications.

Lipton, P. 1991. *Inference to the Best Explanation*, London:Routledge.

Mill, J. S. 1973 *A System of Logic*, Books I-III, Volume VII of *Collected Works*. University of Toronto Press and London: Routledge & Kegan Paul.

Reason, J. T. 1989. The contribution of latent human failures to the breakdown of complex systems. In

Broadbent, D. E., Baddeley, A., and Reason, J. T. eds, *Human Factors in Hazardous Situations*, Oxford U. P., 1989.

Reason, J. T., 1990. *Human Error*, Cambridge U. P.

[RVS-**] Various publications available from <http://www.rvs.uni-bielefeld.de>