University of Bielefeld
Faculty of Technology

RVS
Rechnernetze und Verteilte Systeme

Analysing a Friendly Fire Accident with WBA

The 1994 Operation Provide Comfort Shootdown of two U.S. Army helicopters

Peter B. Ladkin, Jörn Stuphorn, Hannes Fehr, Jan Paller, Jan Sanders
Contact: ladkin@rvs.uni-bielefeld.de

Part 1: Overview of Study

Sources

- U.S. Army Lt. Col. (now Col.) Scott Snook, Friendly Fire, Princeton University Press, 2000
  - Sociological study using thin/thick description techniques and theories of individual, group and organisational behavior
- U.S. DoD Aircraft Accident Investigation Board Report, Executive Summary (WWW source)

Goal of the Study

- Originally: to analyse the incident using WBA and compare with Leveson's STAMP analysis, 20^{th} ISSC, 2002
- Actually: to redo Snook's analysis using WBA techniques
  - To represent Snook's analysis as causal explanation
  - To devise methods to represent social factors and theories effectively within a WB-type analysis
- As the study progressed:
  - To check for consistency and resolve explanatory conflicts
  - To devise methods of composing small WB-graphs
  - To devise meaning-preserving methods of fusing nodes

Ab initio Social Analysis: What is Needed

- Identify crucial behaviors needing explanation
- Identify social theories to explain this behavior
  - Individual, group, organisational
  - In this three-way classification, we already impose a theory
  - Additional theories inside each category
  - The choice of social-explanatory theories appears to be imposed on the incident and thus **a priori**
  - One could imagine developing templates for applying a theory
- ! It follows, then, that a SOL is as „good“ as a Snook !

Snook's Method

- Uses a time line and a „causal map“
  - Causal map (very informal use of Counterfactual Test) „insufficient“
  - Four levels of goings-on
    - Individual, Group, Organisational and Technical levels
  - „Thick description“ technique (cf Vaughan, Challenger)
  - Oriented almost entirely towards the human systems
    - Identifies known social and psychological phenomena
    - Applies theories of individual, group, organisational behavior
  - Devises theory of „practical drift“
  - One-shot deal
    - generalisability unclear (to me at least!)

WBA (Ladkin, Loer)

- Identifies the causal factors involved in an incident
- Uses the Counterfactual Test to determine whether one state or event is a causal factor of another (rigorous semantics)
- Tests for correctness and relative (in)sufficiency
- Methods to represent underdetermination and factors of omission
- Backed by a rigorous formal method for determining correctness and relative sufficiency of a proposed explanation, based on Lamport's reverse-natural-deduction proof schemes and a formal logic EL (time, causal, deontic logic + rules of explanation)
- Claims to be general method for explicating causality in all systems

Outcome of the Study

- WBA study shows
  - Modified WBA (Causal Explanatory Graph) indeed applicable to organise social-organisational-psychological factors
  - Verification and correction ability of semi-formal method (WBA in this case) essential for checking informal sociological work

Part 2: The Incident and Snook's Causal Representation

US Army Black Hawks shot down by USAF F-15 aircraft

- 14 April 1994 shootdown of two UH-60 utility helicopters by USAF F-15 aircraft in Northern Iraq during Operation Provide Comfort
  - Col. Scott Snook, sociological study *Friendly Fire, Princeton U.P. 2000*
  - Leveson et al., STAMP, ISSC20, 2002
  - Ladkin et al., WBA, Bieleschweig 3, 12-13 February 2004

Friendly Fire Incident

- Two UH-60 Black Hawk helicopters, „Eagle Flight“, ferrying high-level personnel to and fro inside the No-Fly Zone (NFZ) of Northern Iraq during Operation Provide Comfort
- Two F-15 aircraft performing the first „clean sweep“ of the NFZ for the day, identify low-flying targets
- Eagle Flight monitored sporadically by AWACS, which had nominal command authority over F-15 engagements, but did not know where Eagle Flight was, nor inform F-15 pilots of its potential presence
- UH-60's visually misidentified and shot down by F-15's

Friendly Fire Incident Analyses

- SECDEF's memo on the AAIB report identifies as „errors, omissions and failures“ that
  - The F-15 pilots misidentified the Black Hawks
  - The AWACS crew failed to intervene
  - Eagle Flight ops were not integrated into the Task Force
  - The Identification Friend or Foe (IFF) systems failed
- Snook: four points here show four types of factor
  - Individual, group, organisational, technical (IGO+T)
- Snook addresses mostly the human IGO domains
  - There appears to be no chance of explanation of the IFF behavior

Snook's „Time Line“

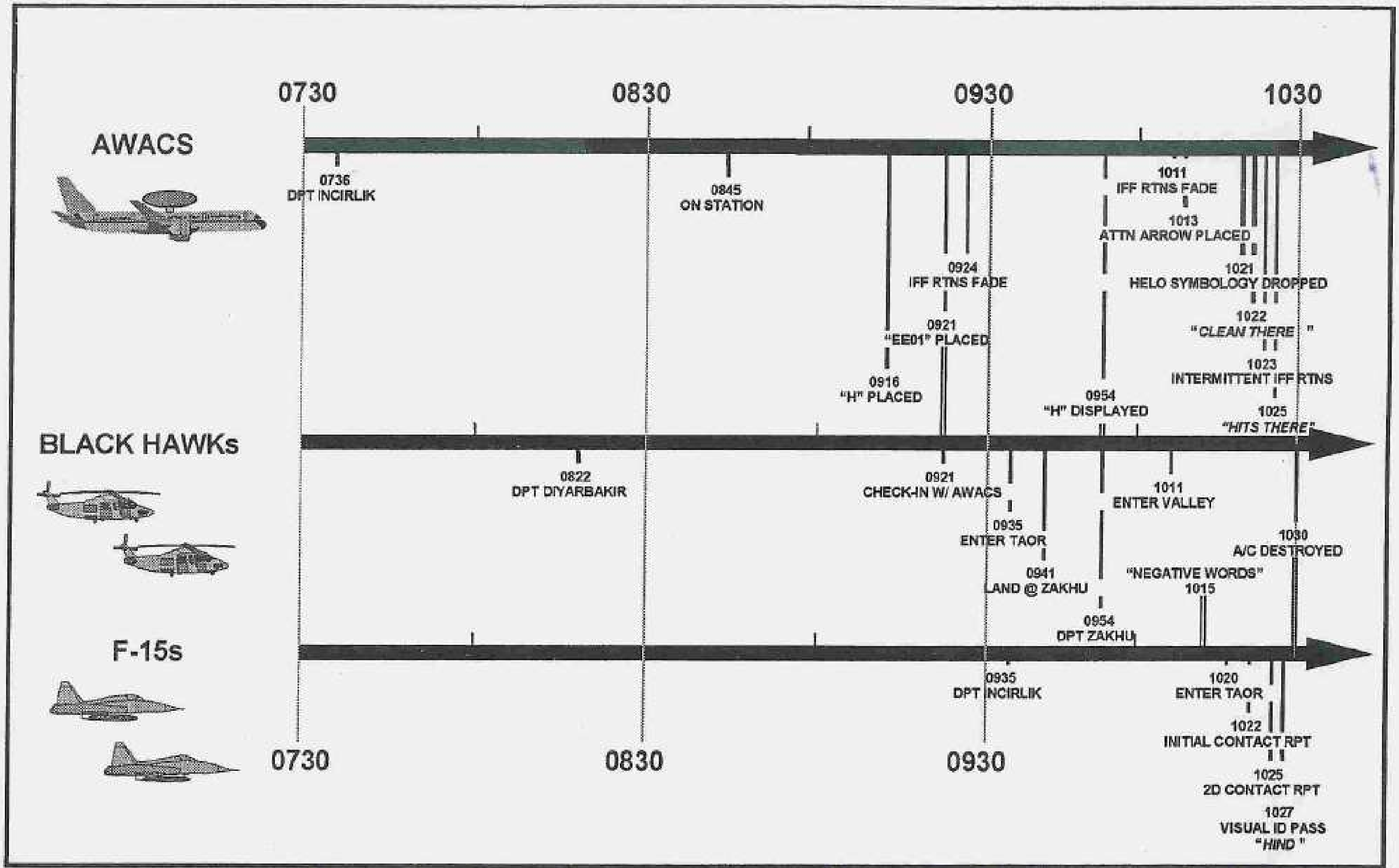


Fig. 1.2. Event Time Line from Scott A. Snook, Friendly Fire, © Princeton U.P. 2000

Snook's Time Line: Critique

- Snook's visual presentation highly unsatisfactory
  - Time Line representation is abstract and confusing
    - It is spatially constrained
    - Due to these spatial constraints, it cannot effectively represent the AAIB Executive Summary events and must abstract
    - Participants in events easily indicated if spatially adjacent, confusingly indicated if spatially separated
    - Simultaneous events and their participants confusingly indicated

Modifying the Representation of the Time Line

- Need way to show arbitrarily detailed (long) time line without spatial constraints
  - Write it vertically rather than horizontally
- Need way to show event participants without visual confusion
  - Participants represented by columns
  - A mark in a column opposite an event represents participation
- Need way to show simultaneous events without confusion
  - One mark on time line for the time
  - Multiple lines by the mark, one for each event
  - Time-separated events separated by a blank line

Our Time Line after Snook

Sheet1

		AWACS	BlHwks	F15
Depart Incirlik	07:36	X		
Depart Dyabirkir	08:22		X	
On Station	08:45	X		
"H" Placed	09:14	X		
Check-in w/ AWACS	09:21	X	X	
"EE01" placed	09:21	X		
IFF returns fade	09:24	X	X	
Enter TAOR	09:35		X	
Depart Incirlik	09:35			X
Land @ Zakhu	09:41		X	
Depart Zakhu	09:54		X	
"H" Displayed	09:54	X		
Enter Valley	10:11		X	
IFF returns fade	10:11	X	X	
Attn Arrow placed	10:13	X		
"Negative Words"	10:15	R		X
Enter TAOR	10:20			X
HELO Symbology dropped	10:21	X		
"Clean there"	10:22	X		R
Initial contact report	10:22	R		X
Intermittent IFF RTNS	10:23	X	X	
"Hits there"	10:25	X		R
Second contact report	10:25	R		X
Visual pass "Hind"	10:27			X
AC destroyed	10:30		X	X

The AAIB Timeline

- The AAIB timeline appears to us to contain 47 significant events
- Snook's timeline includes 25 such events
- It looks as if Snook made a selection
  - The criteria are unspecified
- Our representation of the timeline does not force such a selection

Two-Thirds of AAIB Timeline

Event	*Time (Zulu)*	*F15s (Tiger)*	*Black Hawks (Eagle)*	*AWACS (Cougar)*
Departure Incirlik AB	0436			X
Departure Diyarbakir	0522		X	
On Station	0545			X
„H" displayed on SD radar scope	0612			X
Radio transmission at „Gate"	0621		X	R
Eagle Track annotated „EE01"				X
Eagle land at Zakhu	~0624		X	
Eagle IFF and Radar fade			X	X
Departure Incirlik AB	0635	X		
Tiger IFF Mode IV interogated	0636	X		X
Onroute from Zakhu to Irbil	0654			X
Radio call received			X	R
„EE01" reinitiated				X
„H" regularly displayed	0655		X	X
Check in	0705	X		R
„H" ceases to be displayed	0711		X	X
Eagle enter mountainous terrain	0712			X
Eagle radar and IFF fade			X	X
Symbology continues at last known speed and direction				X
ASO places SD scope in vincinity of *Eagle* last known position	0713			X
Check in with ACE	~0715	X		R
ACE replies „... negative words"		R		X
Radar adjusted to low-velocity detection				X
Enter TAOR	0720	X		
Radio transmission at „Gate"		X		R
„EE01" symbology dropped	0721			X
Report of radar contacts at 40nm	0722	X		R
Reply „clean there"		R		X
IFF response in vicinity of *Tiger's* radar contact	0723		X	X
„H" symbol reappears	0724		X	X
IFF response more frequent	0725		X	X
Radio „Contact"		X		R
Radio „Hits There"		R		X
Enroute controller initiates „Unknown, Pending, Unevaluated" symbol in vicinity of *Eagle* IFF/radar returns	0727			X
Enroute controller attempts IFF interrogation				X
Visual identification with helicopter at 5 NM	~0728	X	X	
Eagle IFF and Radar fade		?	X	?

Snook's „Causal Map“

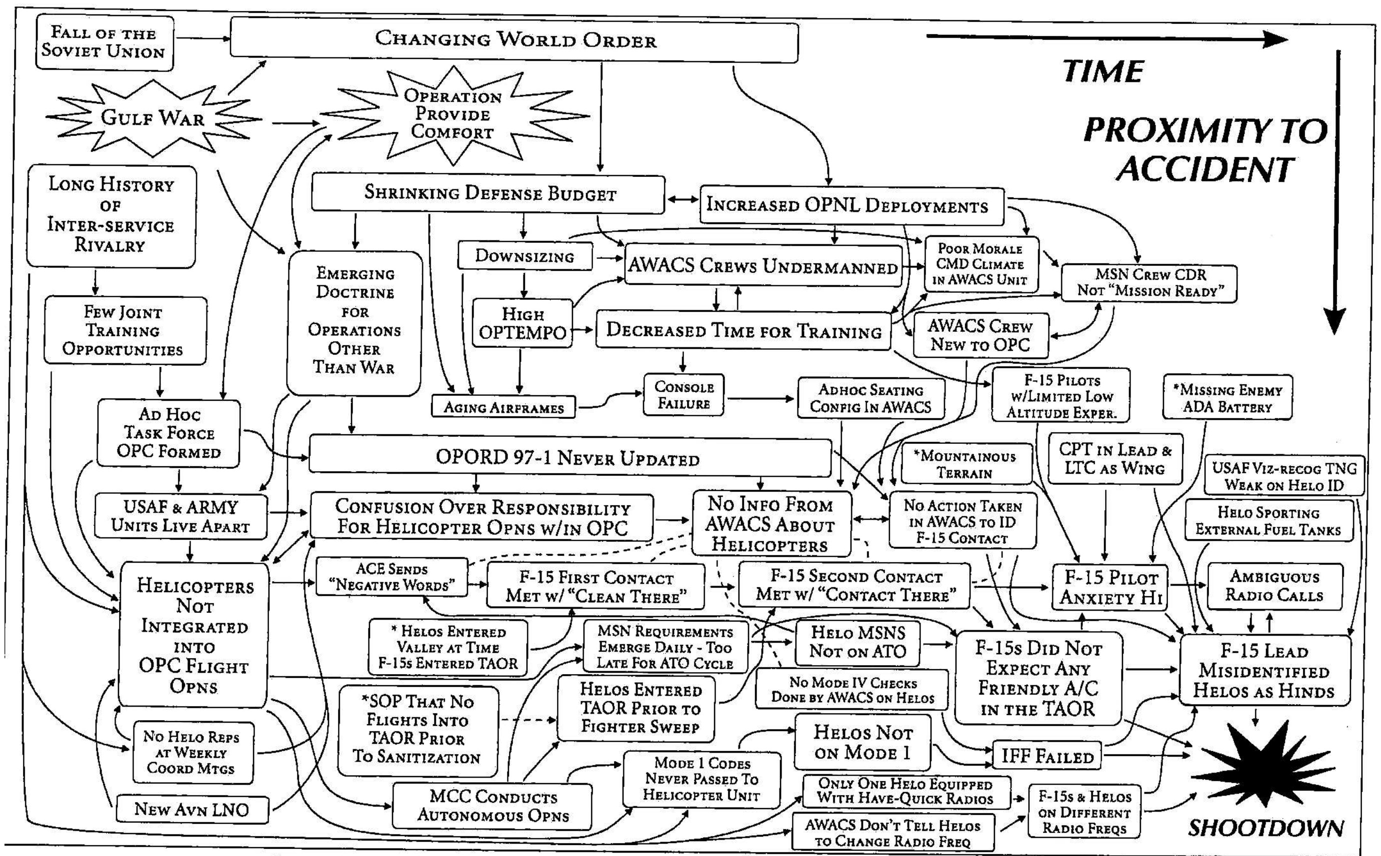
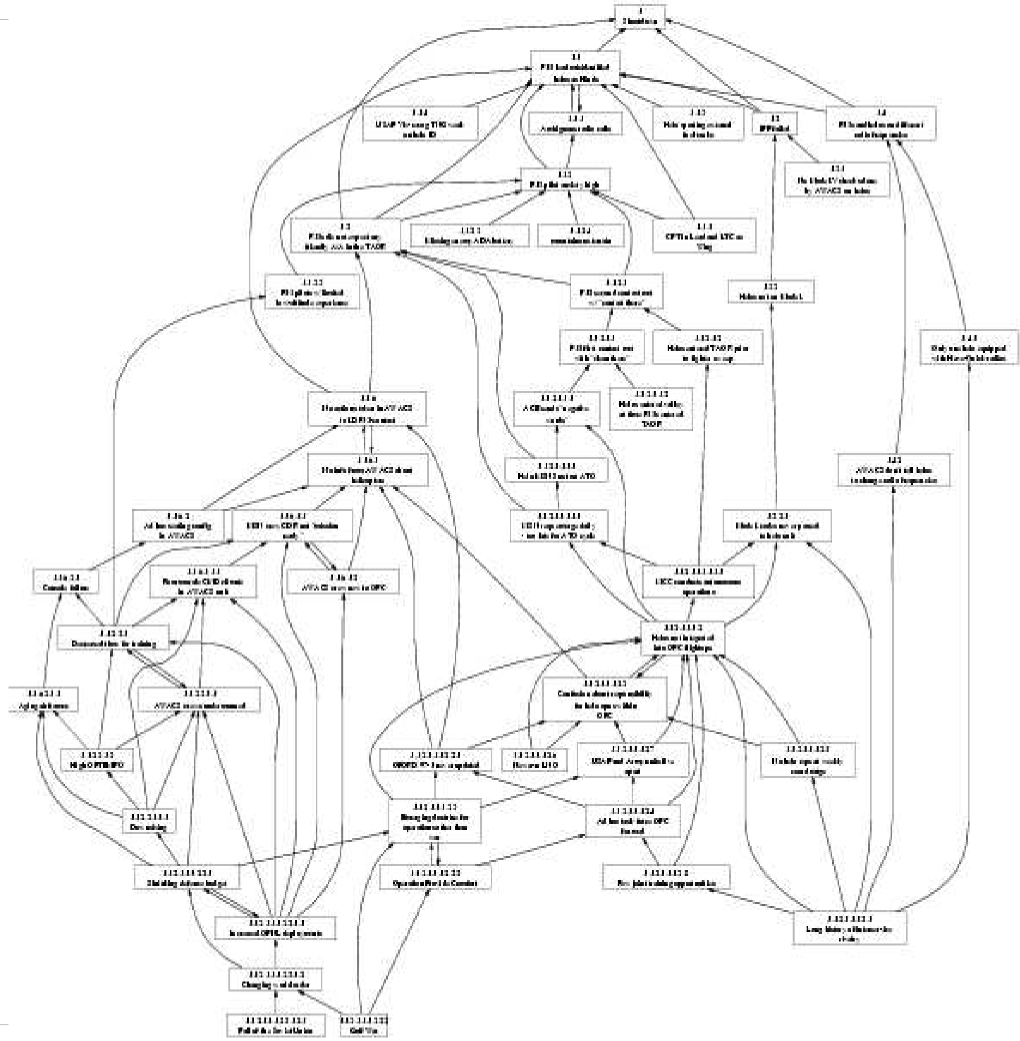


Fig. 1.3. Causal Map from Scott A. Snook, Friendly Fire, © Princeton U.P. 2000

Snook's Causal Map: Critique

- Causal Map is visually confusing
  - It is spatially constrained
  - It does not enable you to see „causal flow“, or (inadmissible) loops
  - It would be hard to „use“, and was hard even to check for correctness

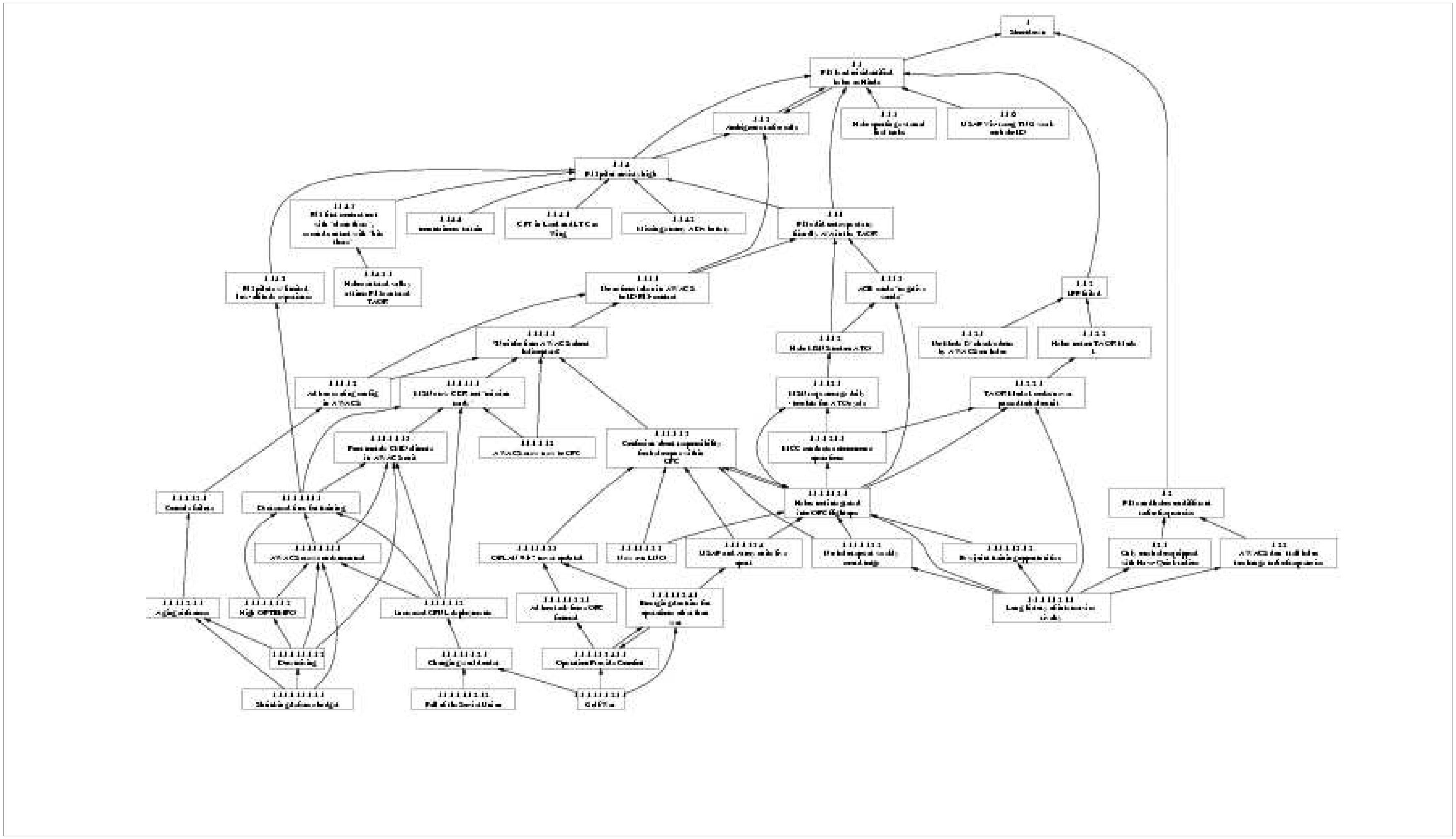
Snook's Causal Map using *ciedit*



Snook's Causal Analysis: Critique

- Different procedures for deriving facts
  - WBA List of Facts derived directly from AAIB Report ExecSumm
  - Snook's list lacks explicit justification; appears to be ad hoc
- Existence of causal loops
  - Seven pairs of phenomena have edges in both directions! This violates the semantics of the Counterfactual Test, which Snook claimed to be using
- Mistakes in application of the Counterfactual Test
  - We found some 69 edges that passed the Counterfactual Test, but also some 25 edges that did not!

Corrected Causal Map



WBA from the AAIB Report Executive Summary

- Our WBG derived from the AAIB report is considered later in Part 5 of this presentation

Critique of Snook's Causal Analysis: Source

- Our critique is presented in full in *Two Causal Analyses of the Operation Provide Comfort Black Hawk Shootdown,* Peter B. Ladkin and Jörn Stuphorn, in series Conferences in Research and Practice of Information Technology (CRPIT), Volume 33, Safety-Critical Systems and Software 2003.

  Available through *crpit.com* , forthcoming in 2004

Part 3: Snook's Evocation of IGO Factors

Summary of Analysis of IGO Factors

- Snook's presentation and analysis is narrative and evocative, with little or no visualisation
- His guiding procedure is curiosity, sociological knowledge, military knowledge, and dissatisfaction with superficial explanatory suggestions
- We summarised Snook's detailed argumentation through highlighting
    - We codified his proposed causes/explanations in mini-WBGs, and then composed the WBGs, as explained in Part 6 of this presentation

Example of the Method: Group Factors

way, there is a lot of nothing going on here—a lot of nothing that needs explaining.

A Weak Team: Overmatched

A focus on AWACS inaction suggests a group level analysis. After all, "it is the team, not the aircraft or the individual pilots, that is at the root of most accidents and incidents" (Hackman, 1993: 49). Shifting attention away from individual pilots to the AWACS crew requires adopting an entirely different analytical perspective—one based on the central premise that groups are important; and should, therefore, be taken seriously (Leavitt, 1975). This chapter takes groups seriously. By tracing the path of our AWACS crew from its formation to its fateful flight, we'll apply what we know about groups in complex organizations to help us address the question: Why didn't the AWACS crew intervene?

The short answer is that the AWACS crew of record flying on 14 April was weak and underdeveloped. It was weak in both an internal, absolute sense and also in an external, relative sense. In an absolute sense, our AWACS mission crew never grew strong enough as a true team to perform beyond a minimum level of proficiency. As long as mission demands remained relatively simple and routine, even our young[6] crew would have performed just fine. Unfortunately, this crew was no match for the unusually demanding set of circumstances they faced on their very first flight together in-country. A weak crew failed to accurately track Eagle Flight helicopters and turned out to be no match for questioning a rapidly developing combat engagement by two fighter pilots. G1

Ultimately, the AWACS crew's ineffective monitoring of Eagle Flight and failure to intervene can be traced back to a fundamental leadership failure. In short, our AWACS crew experienced a very poor launch. Key leaders failed at the critical task of crew formation. An overreliance on organizationally defined positions, standard operating procedures, and interaction rules led to the unquestioned adoption of a priori scripts as shallow functional substitutes for more deeply shared norms.[7] This is not all that surprising, given the Air Force's historical emphasis on *individual* training and qualification. Air Force personnel systems are primarily designed to select, train, and qualify G2

[6] In this sense, I use the word "young" to refer to the collective age or "mission experience age" of the crew. Individual crew members may have been quite experienced; however, as a team, our AWACS crew was still young in the sense that they had not shared any significant work experiences *together* yet as a team.

[7] Bettenhausen and Murnighan (1985) also found that "deliberate groups," whose members invested time up front to negotiate and fine-tune expected norms, performed better when subsequent difficulties emerged than did "impetuous groups," who proceeded quickly and confidently, assuming similar a priori scripts.

individual crew members, not intact teams. For example, after the shootdown, a great deal of time was spent trying to determine if the Mission Crew Commander was technically "mission ready" in accordance with Air Force Regulations. No emphasis was placed on the relative "mission readiness" of the ultimate performing unit—the crew as a real team. As a collection of individuals, they may have been "technically qualified"; as a team, they remained "collectively weak"—weak in an internal, absolute sense. An ad hoc group of individuals thrown together for this particular rotation never really gained a true sense of mutual responsibility and accountability for collective outcomes. In short, the mission crew had not yet developed into a "real team." G3

Comparing them to other mission groups—in an external, relative sense—this failure to develop into a strong team virtually guaranteed AWACS's subordinate position within the larger OPC supra-system, a position that significantly detracted from their ability to control OPC mission aircraft. All complex organizations develop informal status hierarchies. The Air Force is no exception. First, there are two types of people in the Air Force: those that wear wings and those that don't—pilots and all lesser mortals. Second, there is a further distinction among pilots. There are fighter pilots—those steely eyed warriors who fly "fast movers"; and then there are "bus drivers"—lesser beings who drive slow-moving cargo and tanker aircraft. If you aspire to the highest position in the Air Force, if you want to be Chief of Staff, you'd better be a fighter pilot. If fighter pilots sit at the top of the status pyramid, you can imagine where a nonrated (no wings-wearing) air traffic controller sits. Given the structurally privileged position of fighter pilots, even a strong mission crew would have been stretched to intervene in a fighter intercept based on sketchy information. Hence, it follows that a weak crew—something less than a "real team"—operating from a relatively low position within the established social hierarchy, would be doubly handicapped. Under such conditions, crew inaction becomes less mysterious. G4 G5

Katzenbach and Smith (1993) differentiate between "real teams" and other levels of performing groups. They also suggest significant performance implications associated with various degrees of team effectiveness. See Figure 4.1 for a diagram of their "team performance curve" and some useful definitions of notional points along this theoretical continuum.

On the morning of 14 April, for a number of reasons that I address below, our AWACS mission crew fell somewhere down in the performance basement between "working groups" and "potential teams." Though they considered themselves a team, their combined effectiveness was actually worse than what you would predict by simply aggregating individual capabilities. In the language of Katzenbach and Smith (1993: 91), they were a *pseudo-team*—"the weakest of all groups in terms of performance impact. They almost always contribute less to company performance needs than *working* G6

Individual Factors: Summary

- Why were UH-60's misidentified by F-15's?
  - Ambiguous stimulus
  - Powerful set of expectations (no other friendly traffic could be there)
  - „Seeing what one wants to see“
  - (Theory: when **stimulus ambiguous**, we interpret using our **expectations** and **desires**.)

Individual Factors: Detail 1

- No „pilot error" in sense of negligence (Rasmussen, Perrow)
  - Suggests to follow Reason to search for causes in context
  - Deemphasises decision making (Allison, Janis, Vaughan) in favor of inquiring after the construction of meaning (Weick)

Individual Factors: Detail 2

- Ambiguous stimulus
  - Looks at training, cites AAIB work
- Expectations
  - Cites AAIB work to determine what they were
  - Expectations influence perception (Perrow, Bruner, especially Weick)
  - Social interaction of TIGER lead + wingman
- Desires
  - Constructed simply from AAIB and military knowledge

Individual Factors: TIGER Social Interaction

- Social Interaction of TIGER lead + wingman
  - „Reality is in part socially constructed“
  - Weick „sensemaking“: interplay between expectations and understanding during short radio calls
  - Inverted hierarchy: little known
  - TIGER 2
    - „mindless“ (Langer),
    - Slipping into obedient role (Zimbardo, Milgram classics)
  - Bruner: expectations affect „perceptual threshold“, the amount of time and input needed
  - High level of arousal invokes „reflexive“ (overlearned, overtrained, dominant) response (Spence, Weick, English and English, Eysenck)
  - Benefits and liabilities of habitual routines (Gersick and Hackman)

Group Factors: Summary

- Why did AWACS not intervene?
  - (Over)reliance on „organisational shells“ to form AWACS crews
  - Abdication of responsibility, claimed diffuse and ambiguous, explained by „social impact theory“
  - Weak team in unfortunate sequence of events

Group Factors: Detail

- Secretary Perry: multiple „failures to act“
- Snook: „causal significance of nonevents“
- Hackman: team (not individuals or aircraft) at the root of most accidents
- Overreliance on organisationally defined positions, SOPs, interaction rules →adoption of „a priori scripts“
  - (Bettenhausen and Murnighan: „deliberate groups“ perform better than „impetuous groups“)
- Katzenbach and Smith: „real teams“ perform better than others; propose team performance curve

Group Factors: Detail 2

- „Unhealthy organisational soup“: operational history and command climate in „rapidly deteriorating supra-system“
- Organisational shells: Ginnet and Hackman studied aircrew performance and „organisational shells“:
  - Military and service culture
  - TF OPC organisation
  - Crew formation
  - AWACS crew at work
- Provide predefined set of interactions
- Katzenbach and Smith: how „working group“ → „real team“?

Group Factors: Details 3

- AWACS leaders' tendency to rely on shells; inhibits their critical actions
  - Kerr and Jermier: certain individual, task, organisational variables can exert more influence over subordinates than their superior
- Spin-up training ineffective
  - Hackman: training with a full crew is important
- Presence of „shadow crew“
  - Presence of „experts“ may have contributed to confused authority relationships and diffused responsibility
- Plain bad luck:
  - screen inoperative & operators moved; very testing circumstances

Group Factors: Diffusion of Responsibility

- Considered major issue: gets 16 pages
- Latané & Darley: „unresponsive bystander“ experiments
- Latané: Social Impact Theory:
  - Strength, immediacy and number of social forces affects intensity of social impact
  - Principle of divided impact: probability that you will help in an emergency if alone is higher than if you think you are with others
- Brown reformulates:
  - Social definition: see others not responding, redefines the situation
  - Diffusion of responsibility: similar to divided impact

Group Factors: Diffusion of Responsibility 2

- Authority relationships confused
  - Partly due to OPC structure
  - Pilots at center → conflicts of authority
  - Powerful informal hierarchy at odds with demands of formal role requirements necessary for reliable functioning of system
  - Only AAIB cited for this part

Organisational Factors: Summary

- Why was Eagle Flight not integrated into Task Force?
  - Co-commander of Task Force was rated F-16 pilot, flew sweeps; was also regular Eagle Flight passenger, but organisation complexity too great
  - Army and Air Force operations highly separated (history)
  - AWACS, F-15s and UH-60s had different orientations toward goals, time, and interpersonal relationships
  - Mismatch between structure of tasking and structure of tasks
  - Plain bad luck
    „For some 1,109 days, coordination by standardization, plan and mutual adjustment adequately handled the challenge of integration. On 14 April 1994, these mechanisms failed“

Organisational Factors: Details

- Lack of coordination between services
  - Historical, recounted by Snook; also confirmed by Allard
- Issues of differentation of tasks and integration into whole
  - Theory of Lawrence and Lorsch
  - Subunits can differ in
    - Orientation towards goals
    - Orientiation towards time
    - Interpersonal orientation
  - There were „integrative challenges“

Organisational Factors: Details 2

- Interdependence: failures to coordinate
  - Typology of Thompson
    - Pooled interdependence: coordinate mechanism is standardisation
    - Sequential interdependence: coordination mechanism is planning
    - Reciprocal interdependence: coordinate mechanism is mutual adjustment
  - Coordination by standardisation failed
    - IFF frequency use uncoordinated
    - ATO and its relationship with Eagle Flight ops
  - Coordination by plan failed
    - F-15s were unaware of the habitual operations of Eagle Flight
    - Hackman: the „stuff of work“ of TIGER resp. Eagle Flight led each to „lose sight of the larger picture“

Organisational Factors: Details 3

- Interdependence: failures to coordinate (cont'd)
  - Coordination by mutual adjustment failed
    - Different radios, and „Min Comm“ behavior
    - All Snook, from AAIB
- „Fallacy of Centrality“
  - Commander Pilkington was F-16 pilot, conducted sweeps, also flew on Eagle Flight
  - In his world, everything seemed to be coordinated
    - F-16 squadrons knew about Eagle Flight
  - Snook cites Westrum, also Weick

Cross-Levels Account

- Theory of Practical Drift
  - General characteristics of organisational engineering
  - Loose/tight coupling, and Rule-based/Task-based logics of action
  - Introduced by Snook

Cross- Levels Account: Theory of Practical Drift

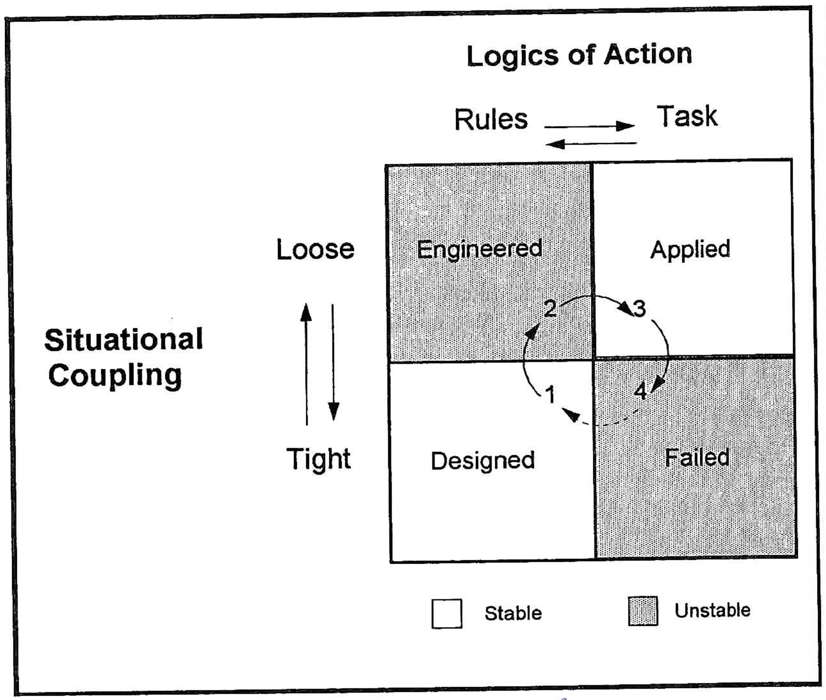


Fig. 6.1. Theoretical Matrix

from Scott A. Snook, Friendly Fire, © Princeton U.P. 2000

Part 4: Comparison of Snook with STAMP Analysis

Friendly Fire: STAMP Analysis

- Uses (complex) Task Force organisational chart from Snook
- Analyses the absence or inappropriateness of implied F&C functions (according to Leveson's schema)
- Finds lots of missing stuff

However....

- Doesn't appear to build a timeline (unclear whether one is required, or what its status would be)
- Doesn't appear to need a „causal map“

Friendly Fire: STAMP Analysis

- Does not explain **why** the control functions were missing
- Fails to identify certain IGO phenomena such as this example
  - Different subsystems whose procedures or worldviews are in tension work out ways of coping at their mutual boundaries (Snook; application of Thompson's *reciprocal interdependence*)
  - „Coping“ with mutually inconsistent procedures means that someone's rules will be violated
  - „Coping“ with incomplete procedures means that existing procedures do not cover the case
- Snook's Practical Drift tries to identify and explain these phenomena

Crude Comparison of Snook and STAMP on FF

- Snook's description is richer, thereby more compelling
- Snook has a theory of how things developed, based on known/socially-theorised IGO phenomena
- However, there is no check whether his theory accounts for all factors, or whether it applies more generally
  - Its plausibility is intuitive and sui generis

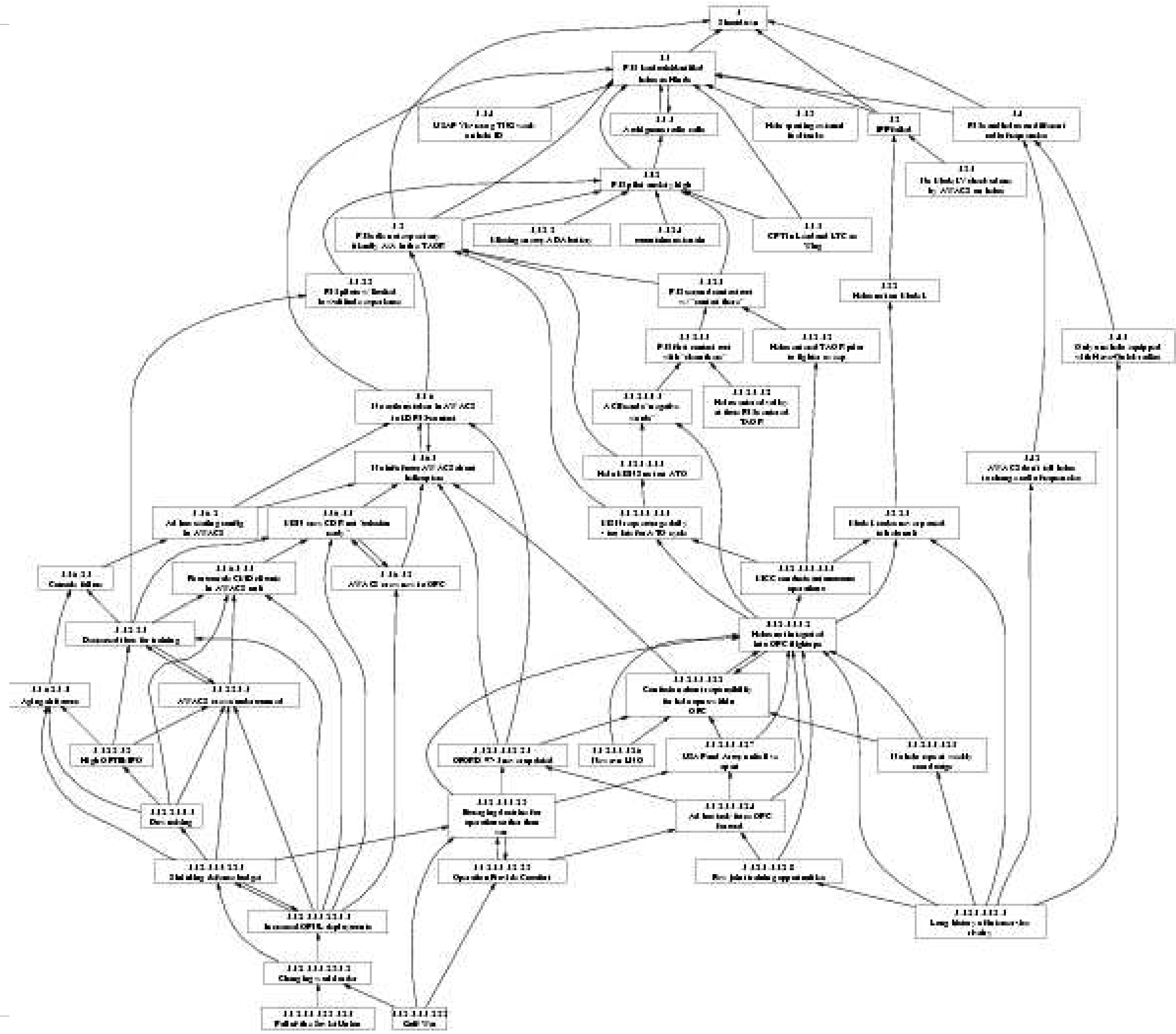
- STAMP works on a „checklist“; it is claimed to be general; application to this example is helped by an intuitive reading of military hierarchy as control structure

Part 5: WBA from AAIB Report Executive Summary

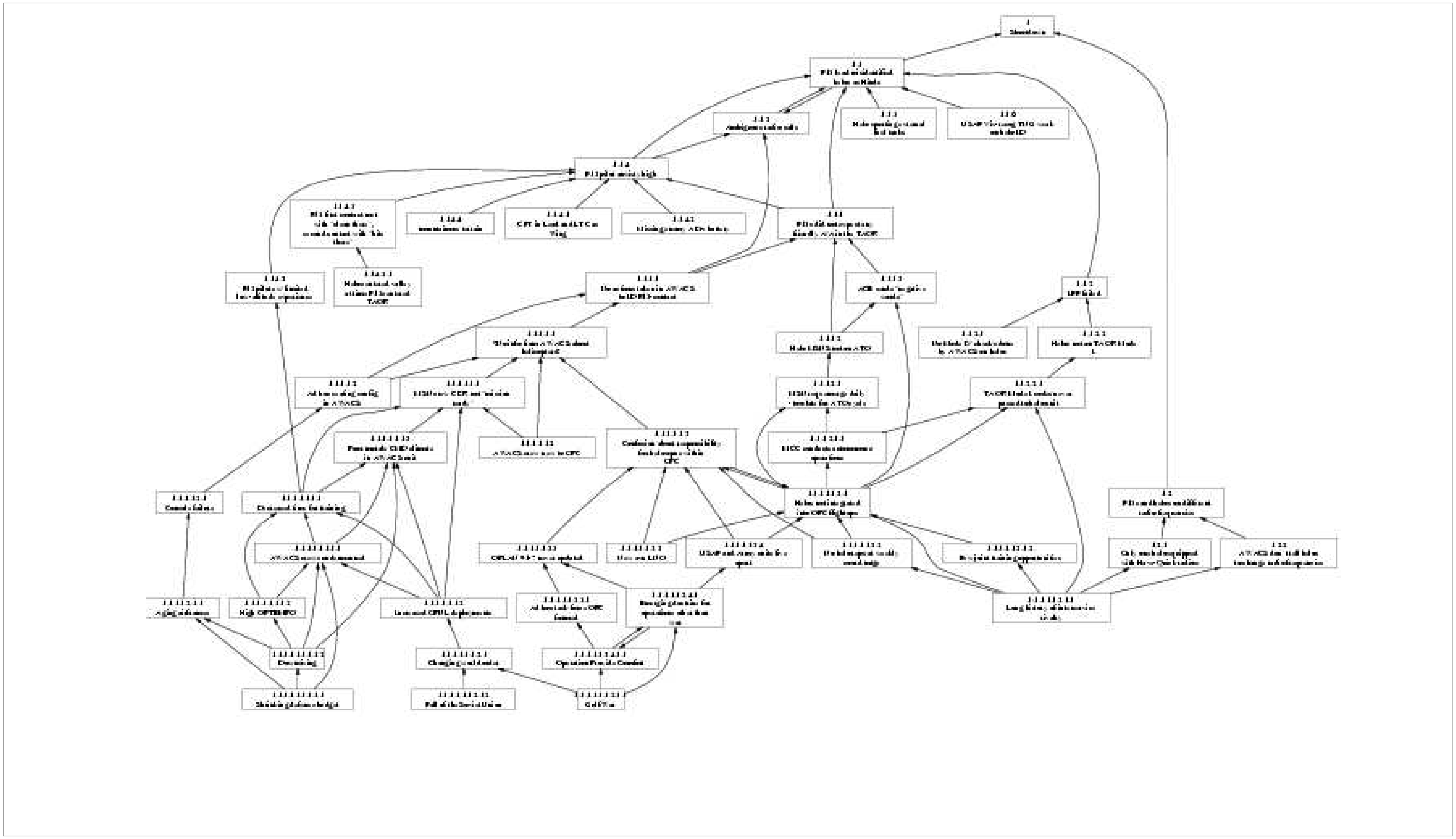
WBA of Friendly Fire from AAIB Report Summary

- We started from USAF AAIB report Executive Summary
- „List of Facts“ different from those of Snook (who does not say how he derived his)
- Causal Map includes 7 double edges (mutual causal influence): but loops are impossible with counterfactual semantics (which Snook claims to use)!
- There are other causal links in the Causal Map which do not pass the Counterfactual Test (so must be mistakes)
- Snook's representation of Causal Map is visually obscure

Snook's Causal Map again



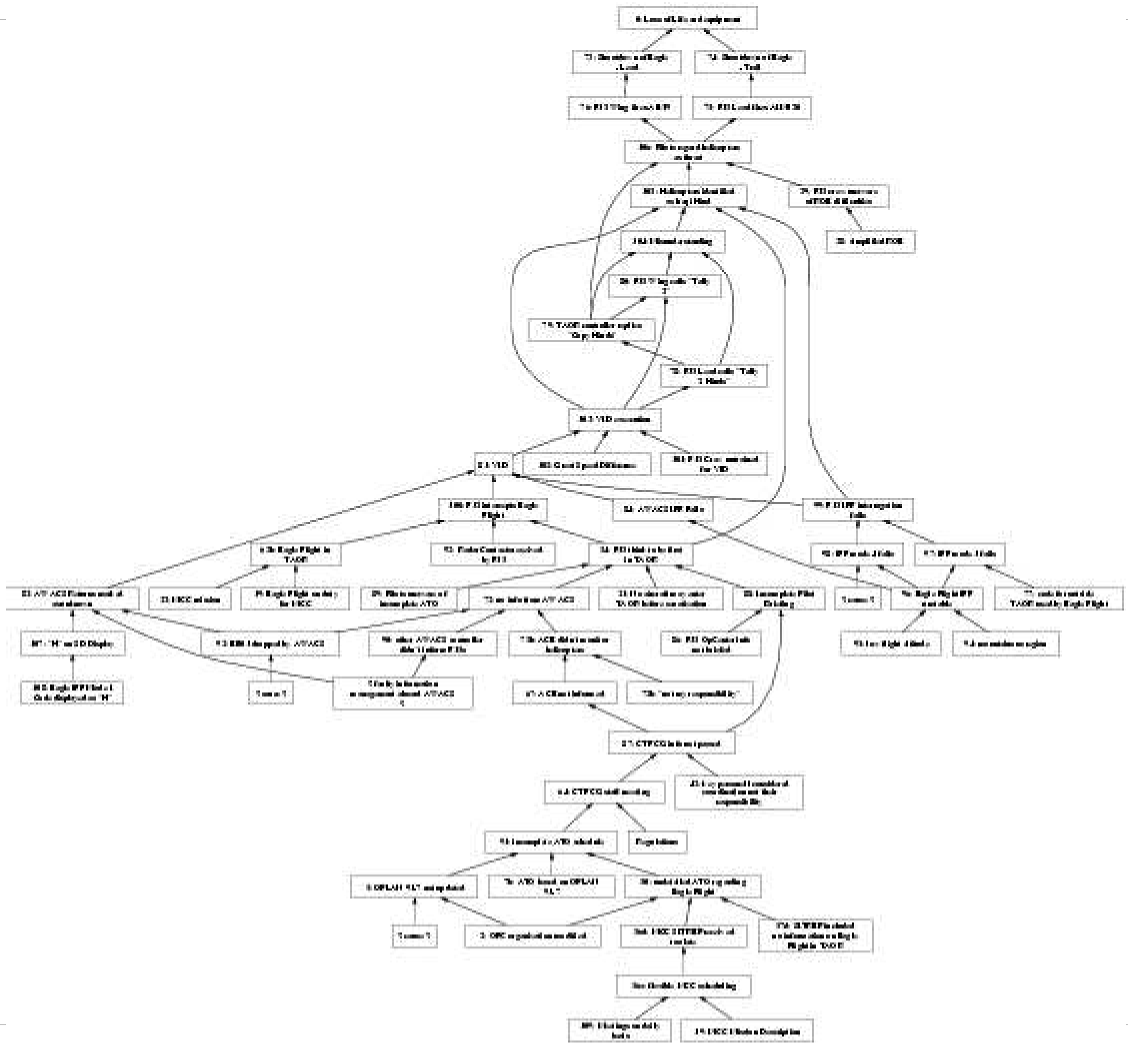
Corrected Causal Map



University of Bielefeld
Faculty of Technology

RVS
Rechnernetze und Verteilte Systeme

WBG from AAIB Report Summary



Part 6: WB-Graph Manipulations

How We Derived Lists of Facts from Snook

way, there is a lot of nothing going on here—a lot of nothing that needs explaining.

A Weak Team: Overmatched

A focus on AWACS inaction suggests a group level analysis. After all, "it is the team, not the aircraft or the individual pilots, that is at the root of most accidents and incidents" (Hackman, 1993: 49). Shifting attention away from individual pilots to the AWACS crew requires adopting an entirely different analytical perspective—one based on the central premise that groups are important; and should, therefore, be taken seriously (Leavitt, 1975). This chapter takes groups seriously. By tracing the path of our AWACS crew from its formation to its fateful flight, we'll apply what we know about groups in complex organizations to help us address the question: Why didn't the AWACS crew intervene?

The short answer is that the AWACS crew of record flying on 14 April was weak and underdeveloped. It was weak in both an internal, absolute sense and also in an external, relative sense. In an absolute sense, our AWACS mission crew never grew strong enough as a true team to perform beyond a minimum level of proficiency. As long as mission demands remained relatively simple and routine, even our young[6] crew would have performed just fine. Unfortunately, this crew was no match for the unusually demanding set of circumstances they faced on their very first flight together in-country. A weak crew failed to accurately track Eagle Flight helicopters and turned out to be no match for questioning a rapidly developing combat engagement by two fighter pilots. G1

Ultimately, the AWACS crew's ineffective monitoring of Eagle Flight and failure to intervene can be traced back to a fundamental leadership failure. In short, our AWACS crew experienced a very poor launch. Key leaders failed at the critical task of crew formation. An overreliance on organizationally defined positions, standard operating procedures, and interaction rules led to the unquestioned adoption of a priori scripts as shallow functional substitutes for more deeply shared norms.[7] This is not all that surprising, given the Air Force's historical emphasis on *individual* training and qualification. Air Force personnel systems are primarily designed to select, train, and qualify G2

[6] In this sense, I use the word "young" to refer to the collective age or "mission experience age" of the crew. Individual crew members may have been quite experienced; however, as a team, our AWACS crew was still young in the sense that they had not shared any significant work experiences *together* yet as a team.

[7] Bettenhausen and Murnighan (1985) also found that "deliberate groups," whose members invested time up front to negotiate and fine-tune expected norms, performed better when subsequent difficulties emerged than did "impetuous groups," who proceeded quickly and confidently, assuming similar a priori scripts.

individual crew members, not intact teams. For example, after the shootdown, a great deal of time was spent trying to determine if the Mission Crew Commander was technically "mission ready" in accordance with Air Force Regulations. No emphasis was placed on the relative "mission readiness" of the ultimate performing unit—the crew as a real team. As a collection of individuals, they may have been "technically qualified"; as a team, they remained "collectively weak"—weak in an internal, absolute sense. An ad hoc group of individuals thrown together for this particular rotation never really gained a true sense of mutual responsibility and accountability for collective outcomes. In short, the mission crew had not yet developed into a "real team." G3

Comparing them to other mission groups—in an external, relative sense—this failure to develop into a strong team virtually guaranteed AWACS's subordinate position within the larger OPC supra-system, a position that significantly detracted from their ability to control OPC mission aircraft. All complex organizations develop informal status hierarchies. The Air Force is no exception. First, there are two types of people in the Air Force: those that wear wings and those that don't—pilots and all lesser mortals. Second, there is a further distinction among pilots. There are fighter pilots—those steely eyed warriors who fly "fast movers"; and then there are "bus drivers"—lesser beings who drive slow-moving cargo and tanker aircraft. If you aspire to the highest position in the Air Force, if you want to be Chief of Staff, you'd better be a fighter pilot. If fighter pilots sit at the top of the status pyramid, you can imagine where a nonrated (no wings-wearing) air traffic controller sits. Given the structurally privileged position of fighter pilots, even a strong mission crew would have been stretched to intervene in a fighter intercept based on sketchy information. Hence, it follows that a weak crew—something less than a "real team"—operating from a relatively low position within the established social hierarchy, would be doubly handicapped. Under such conditions, crew inaction becomes less mysterious. G4 G5

Katzenbach and Smith (1993) differentiate between "real teams" and other levels of performing groups. They also suggest significant performance implications associated with various degrees of team effectiveness. See Figure 4.1 for a diagram of their "team performance curve" and some useful definitions of notional points along this theoretical continuum.

On the morning of 14 April, for a number of reasons that I address below, our AWACS mission crew fell somewhere down in the performance basement between "working groups" and "potential teams." Though they considered themselves a team, their combined effectiveness was actually worse than what you would predict by simply aggregating individual capabilities. In the language of Katzenbach and Smith (1993: 91), they were a *pseudo-team*—"the weakest of all groups in terms of performance impact. They almost always contribute less to company performance needs than *working* G6

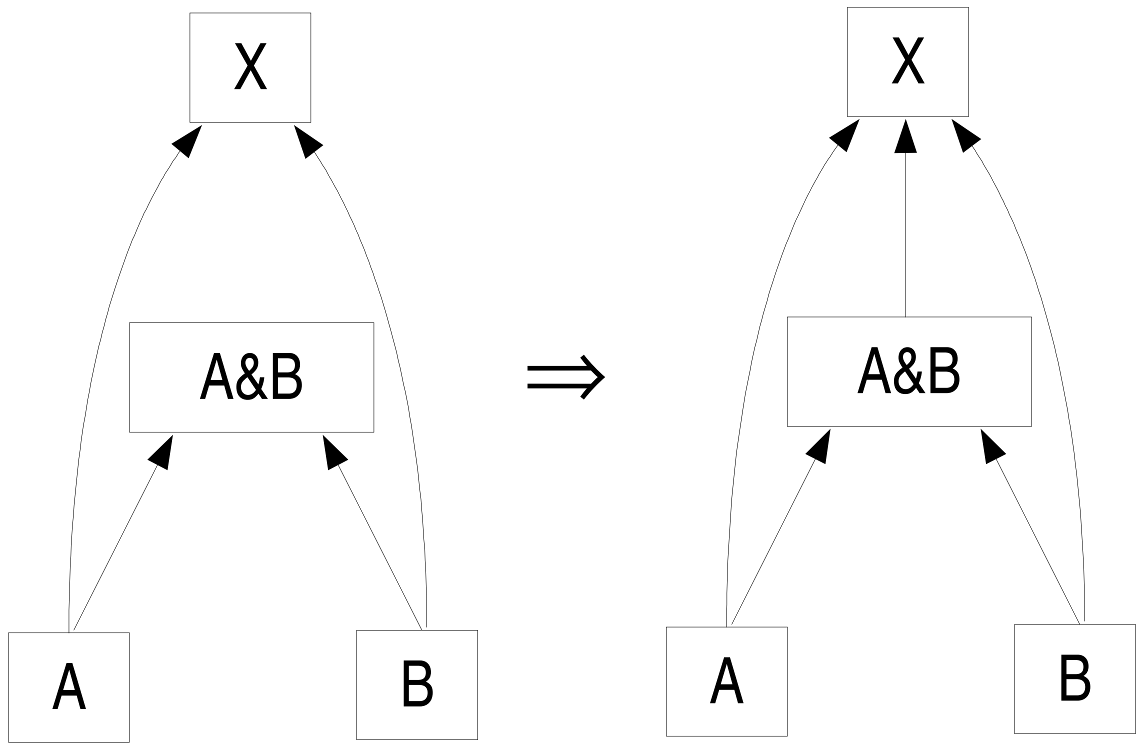
Method

- About 40pp per factor set
- We made mini-graphs from separate underlined „causes“
- We needed to compose the graphs
- Many nodes were equivalent (could be considered alternative ways of saying the same thing), and some equivalent nodes appeared in many mini-graphs

Formal Operations on Graphs

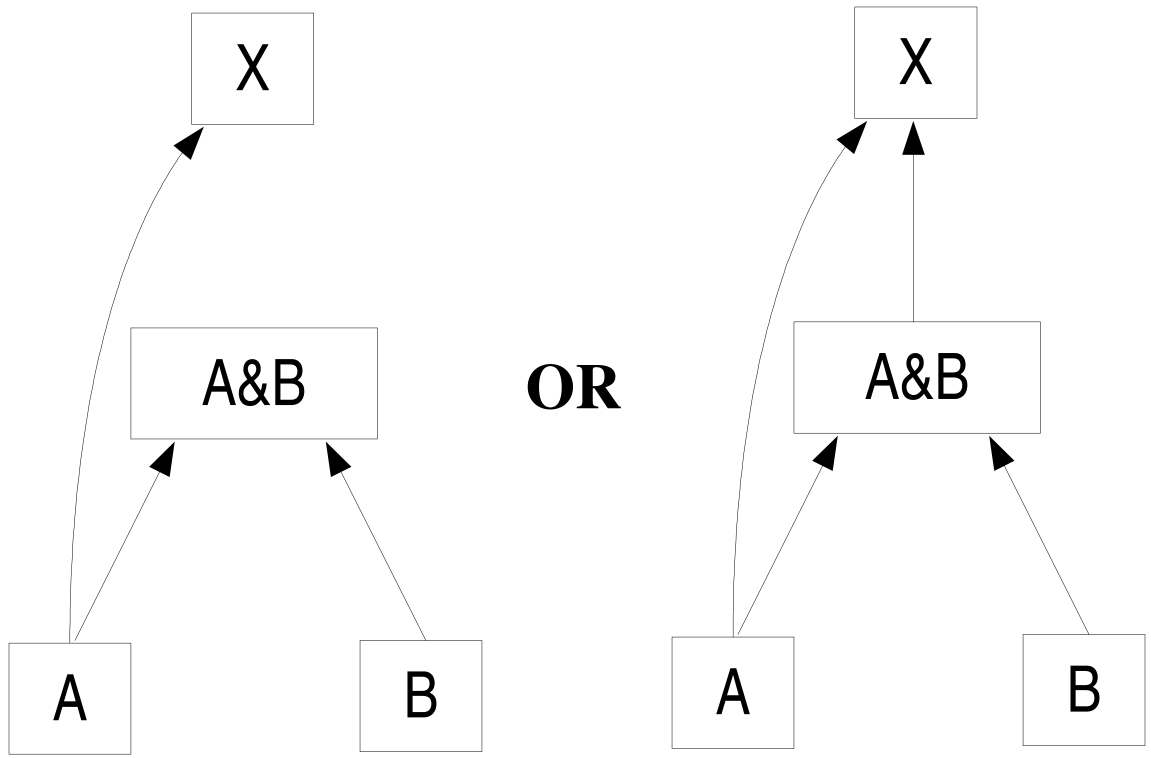
- We needed a fuse-nodes operation
  - In the following, we omit the detailed semantic argumentation from possible worlds needed to justify the operations we propose
- We needed a compose-graphs operation
  - This should be formal, as simple and general as possible, but yet yield the smallest WBG containing given WBGs *A* and *B*
  - We can denote the result of this operation as the *Counterfactual Closure* of *G1* and *G2*, *CfCl(G1,G2)*

Fusing Nodes 1



When A and B are both NCFs of X, then so is A&B

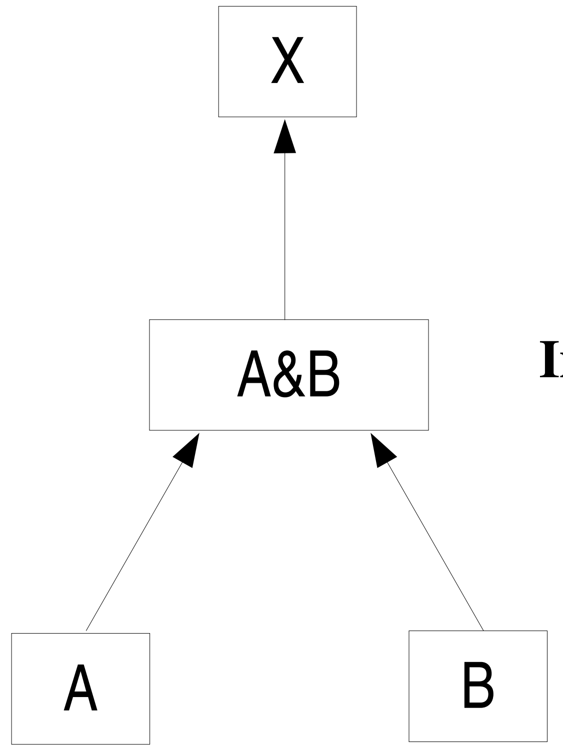
Fusing Nodes 2



If only one of A or B is an NCF of X, then A&B can be an NCF of X or not

Both situations are possible

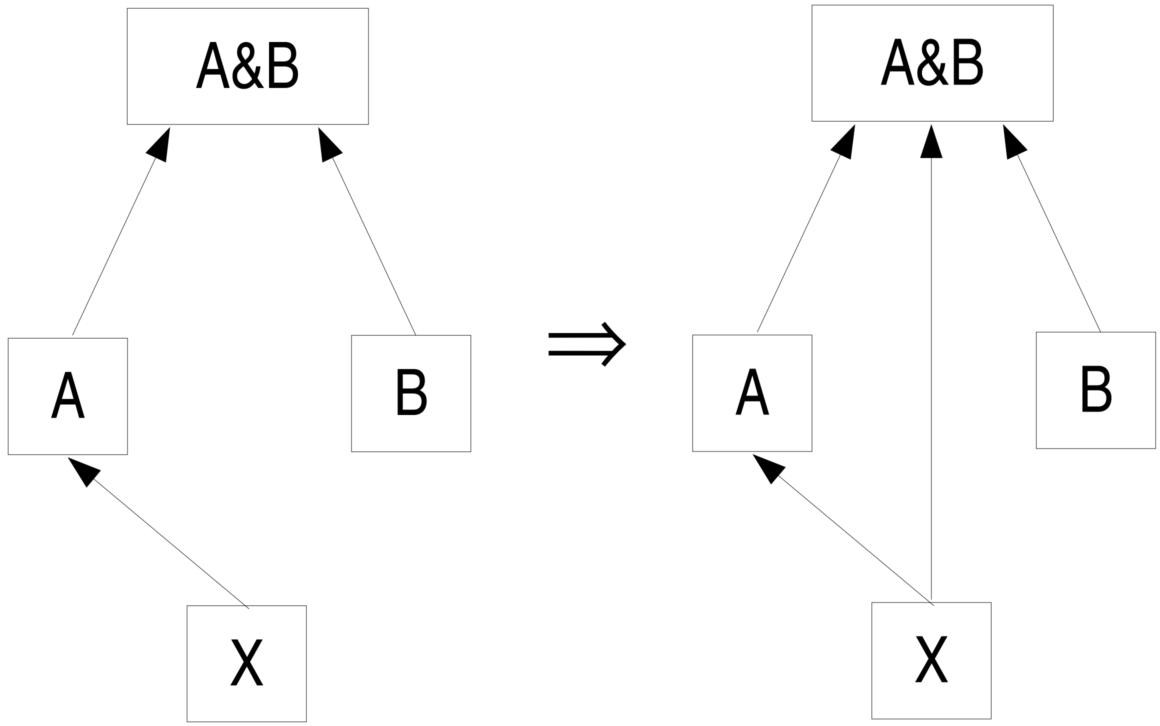
Fusing Nodes 3



Impossible!

It is however not possible for A&B to be an NCF of X without either A or B (or both) being an NCF of X

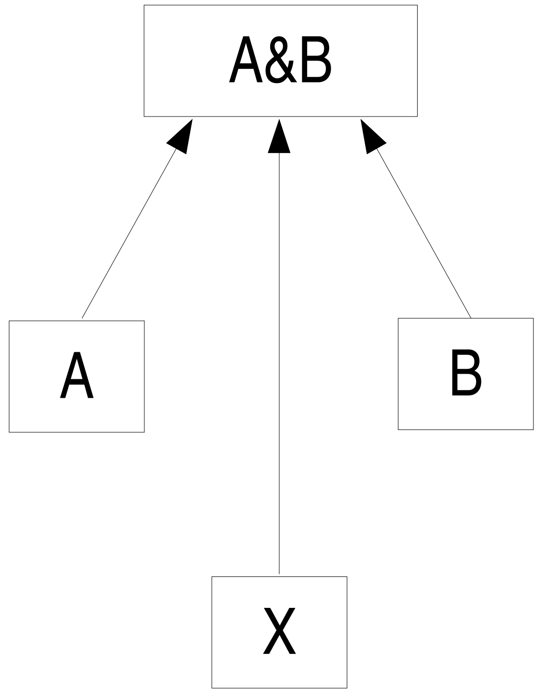
Fusing Nodes 4



Any NCF of A is an NCF of A&B

Any NCF of B is an NCF of A&B

Fusing Nodes 5



Possible!

A node may be an NCF of A&B
even though it is neither an NCF of A nor of B

Fusing Nodes 6

- In some cases, WBG edges can be automatically connected to/from A&B
- There is no obvious algorithm for *CfCl(A,B)* which avoids use of the Counterfactual Test on certain nodes of the composition, but one can maybe minimise its application
- However, when fusing equivalent nodes (A ⇔ B), the rules become simple:
  - All original out-edges are out-edges of the fused node
  - All original in-edges are in-edges of the fused node
  - Watch for inconsistencies (X an NCF of A in one version only!)

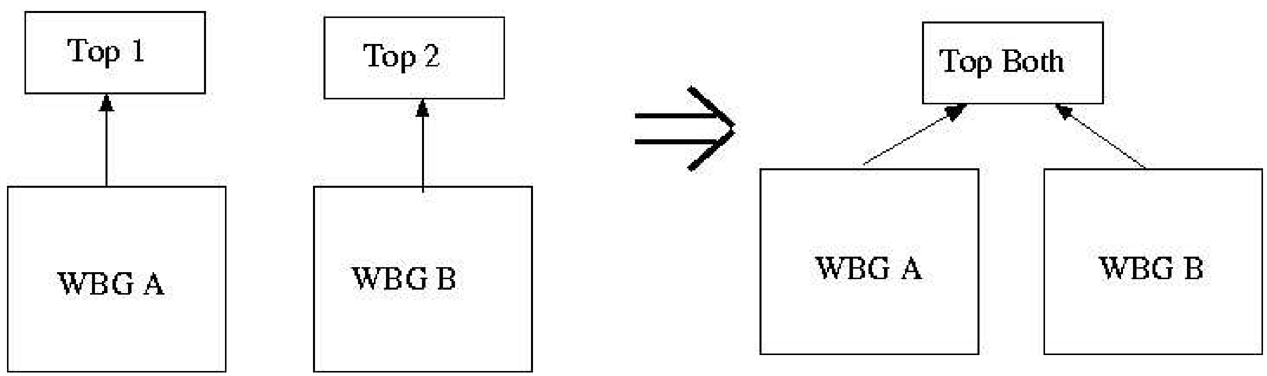
WBG Manipulations

- The Group Factors graph-build was the first we attempted
  - 84 facts from Snook (in final version)
  - 1 fact we added ourselves
  - 9 mini-graphs with 10, 11, 7, 5, 5, 9, 5, 5 and 27 nodes
  - Fusion applied to identical nodes, also to some separate nodes
  - Current state: 3 graphs with 10, 9, 55 nodes, the latter after many fusion & composition operations
  - Fact numbers from List of Facts are essential

Notation for WBG Manipulations

- Graphs *G1* and *G2* may be disjointly composed by adding a „place-holder“ top node and including *G1* and *G2* beneath, eliminating any place-holder top nodes in *G1* and *G2*
  - We denote the disjointly-composed graph by *G1 + G2*
- Then it must be considered through the Counterfactual Test whether any edges should be added between nodes of *G1* and *G2* and vice versa

Disjoint Composition of WBGs



From Disjoint Composition to Counterfactual Closure

- One may fuse identical nodes using the rules given earlier
- Say two nodes x and y have an edge between them in *G1*, but the identical nodes do not have an edge between them in *G2*
  - This contradiction must be resolved manually through use of the Counterfactual Test
- One must also consider applying the Counterfactual Test between nodes of *G1* and of *G2*, and vice versa
- We can denote the result of these operations as the *Counterfactual Closure* of *G1* and *G2, CfCl(G1,G2)*

Notation for Fusing Nodes

- *U(x,y)* to indicate that nodes *x*, *y* have been fused, with the appropriate mandatory edge operations
  - If x occurs in both graphs, we denote the fused node as *U(G1.x, G2.x)*
  - Nodes x and y may have different labels. In this case we say U(x,y)
    - Node (14: Downsizing) in G8 and node (26: Downsizing) in G9 are arguably identical. They become U(14,26) in the composed graph of G8 and G9
    - (16: Shrinking defence budgets) and (24: Shrinking defence budgets) become U(16,24)
    - (17: Increased OPNL deployments) and (23: Increased OPNL deployments) become U(17, 23)
  - One may fuse non-identical nodes

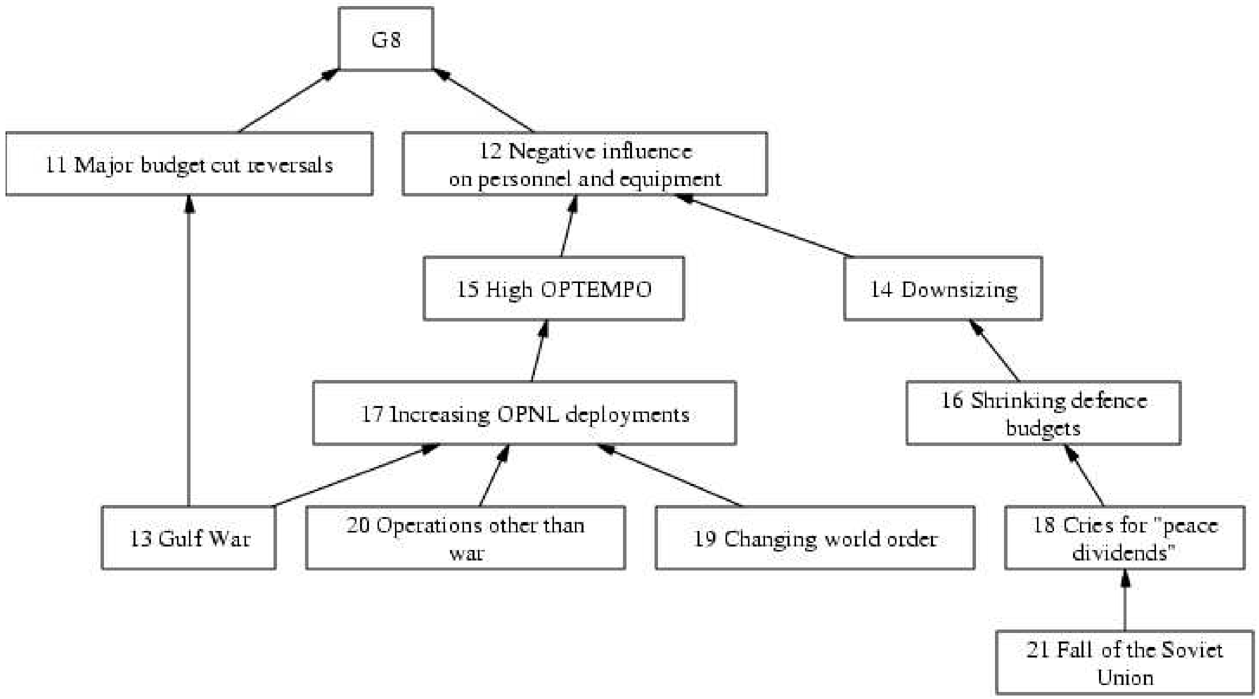
Resolving Edges

- A fused node *U(x,y)* must be checked for additional in- and out-edges in certain cases
- An added edge from *x* to *z* is indicated + [x →z]
- A deleted edge is indicated – [x →z]

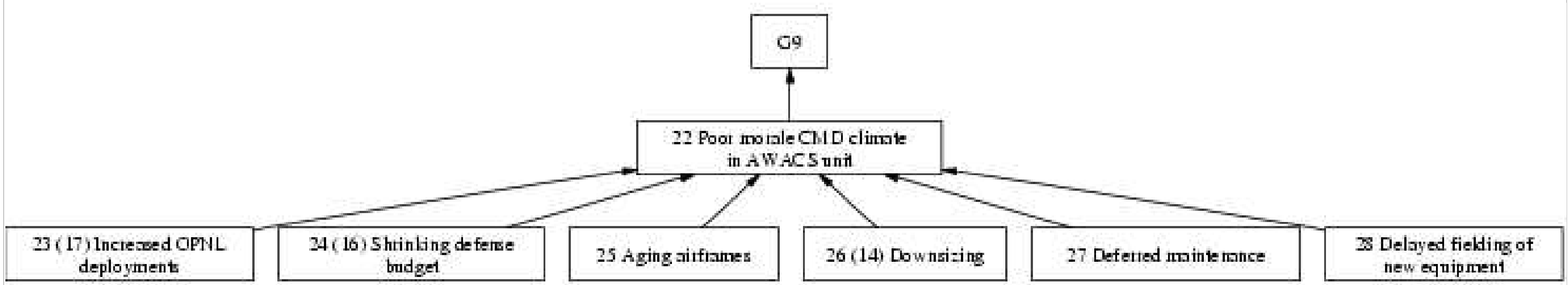
Example of Graph Composition

- One hopes that node fusion is a commutative operation, so one may denote it also by +
  - G8-9 = CfCl(G8,G9)
    = (G8 + G9) + U(14,26) + U(16,24) + U(17,23) + [22 → 12]

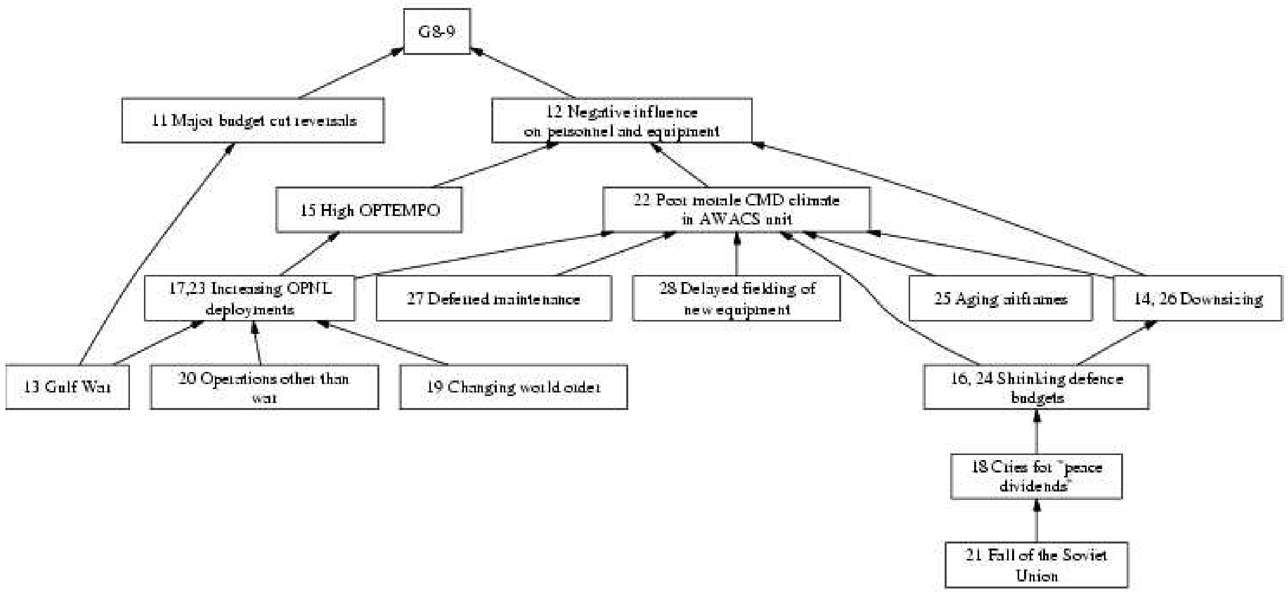
Example: G8


G8
11 Major budget cut reversals
12 Negative influence on personnel and equipment
15 High OPTEMPO
14 Downsizing
17 Increasing OPNL deployments
16 Shrinking defence budgets
13 Gulf War
20 Operations other than war
19 Changing world order
18 Cries for "peace dividends"
21 Fall of the Soviet Union

Example (cont'd): G9



Example (cont'd): G8-9 = CfCl(G8,G9)



Checking Correctness of Composed Graphs

- When a composed graph *F(G1, G2)* has been created from graphs *G1* and *G2*, one can project *F(G1,G2)* on to the node subset *Nodes(G1)* and check that the result is equal to *G1*
  - *Proj(F(G1,G2),G1) = G1*
- Similarly for *G2*
  - *Proj(F(G1,G2),G2) = G2*
- Currently, we do this in *ciedit* by hand
  - It could – will – be automated

Example (cont'd): Checking

- *Proj(G8-9,Nodes(G8)) = G8*
- *Proj(G8-9,Nodes(G9)) = G9*

- Yes, in this case

Part 7: Including Sociological and Other Theories in WBA

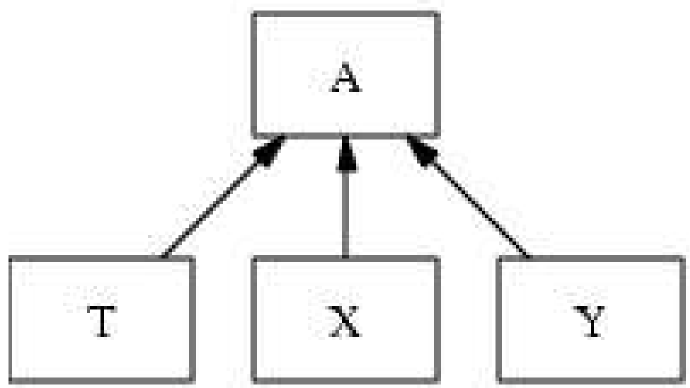
Hempelian Deductive-Nomological Interpretation

- A scientific theory T is a collection of axioms
- A phenomenon A is explained by Theory T under the observations X,Y, according to the deductive-nomological interpretation of Hempel, if and only if
    - A is a logical consequence of T,X,Y
- If A is a logical consequence of T,X,Y, and this is the only explanation offered, then we can argue that
    - T is true, and X, Y are true, and A is true
    - Were T not to be true, then A would not have been true
    - Were X and Y not to have been true, A would not have been true

Deductive-Nomological Explanation (cont'd)

- Remember that the Counterfactual Test is satisfied by logical consequence (that is, its converse!)
- Suppose it is the case that
  - T is true, and X, Y are true, and A is true
  - Were T not to be true, then A would not have been true
  - Were X and Y not to have been true, A would not have been true
- Then T satisfies the semantic condition to be an NCF of A
- X and Y satisfy the semantic condition to be an NCF of A
- We can represent a theory and its premises used as an explanation as follows

Deductive-Nomological Explanations in a WBG



Some Caveats on D-N Interpretation

- Social explanations are often overloaded
  - Many people have theories that explain certain phenomena, and they may use common premises
  - Snook cites multiple social-explanatory sources for a phenomenon
  - In this case, the strict D-N counterfactual is not fulfilled, for the usual reasons of overloading that cause problems for the Counterfactual Test
  - The theories are being proffered as explanations
  - Maybe the arrows in such a graph are better interpreted as causal-explanatory factors and not strictly as NCFs
  - Maybe we should call the result a *Causal-Explanatory Graph*, *CEG*

Causal-Explanatory Graphs

- Causal explanations are arguably transitive, which the NCF relation is not
    - That is, if A is a causal explanation of B, and B is a causal explanation of C, then it follows that A is a causal explanation of C
    - Whereas if X is an NCF of Y, and Y an NCF of Z, it does not follow that X is an NCF of Z (although it may be)
- If so, maybe one should take a CEG to be the transitive closure of a WBG, or of a WBG-with-D-N-explanations
- (We have not taken the transitive closure of the WBG-with-D-Ns that we generated in the reproduction of Snook's analysis)

Part 8: Conclusions

Conclusions

- Method brings you much
- Checking your work is very important
- Devising ways to check your work is very important
- Social-factor theories can find their place in WBA using Hempelian deductive-nomological interpretation
- Causal-explanatory graphs (CEGs) may be more useful for social-scientific factors than pure WBGs
- Interpreting social factors does require one to impose a theory a priori

Conclusions (cont'd)

- Detailed work reproducing social explanations as CEGs is complex
  - Many nodes
  - Many connections to consider
  - Few algorithms, although those that exist help a lot
  - (Semi-)automated tools would be a very great help

Resource Information

- We completed first version of group-factor causal-explanatory graph (CEG) in some 30-40 person-hours
- Organisational-factor CEG equivalent work, but expect 20ph
- Individual-factor CEG took 20+ ph
- Technical-factor CEG is trivial (couple of minutes)
- Fusing is non-trivial
  - But the theory is there; only the tools are lacking
- Factoring the complete CEG into components likely to be non-trivial undertaking
  - Factoring theory urgently needed!

Anticipated (Mature) Effort

- Meetings
  - 4 x 1.5 hours per social-factor set
  - Four factor sets: Technical + 3 social (Snook + also SOL)
    - Individual
    - Group
    - Organisational
  - Each factor set ~40pp of filtered, argued evidence
  - Derived ~100 facts per List of Facts per factor set
  - ~10-15% „summary“ facts introduced, rest directly from Snook

Thanks for listening!