

Confronting context effects in intelligence analysis: How can mathematics help?

Keith Devlin*

July 15, 2005.

INCOMPLETE DRAFT, UNDER DEVELOPMENT

USE CAUTION WHEN OPENING AS CONTENTS MAY
HAVE SHIFTED SINCE YOU LAST SAW THEM

Abstract

We use the interpretation of a key piece of information in the 2002 U.S. decision to invade Iraq to motivate a study of the role of context in decision making. Although our primary focus is intelligence analysis, our study seeks to adopt a mathematical approach. This should make it applicable in a wide variety of application domains. In particular, the level of abstraction at which we work enables us to make useful comparisons between data gathering for intelligence analysis and Internet commerce, both of which depend on making crucial estimates of trust.

The initial examination of the way context influences reasoning and decision making, together with the results of studies carried out by others (some of which we mention), leads us to conclude that the role played by mathematics in improving intelligence analysis will necessarily be different from the role it plays in engineering or the natural sciences. Whereas some aspects of our study may find their way into the design of computer support tools for intelligence analysis — for instance the logical formalism we described in our earlier paper [16] — we believe that the most significant benefits to intelligence analysis from this work are likely to be:

- better appreciation of the way context influences reasoning and decision making;
- sharper insight into the human problems inherent in intelligence analysis;
- improved analysis protocols to guide the intelligence analysis community.

In this respect, our study may be viewed as a mathematical analog of Richards J Heuer's *Psychology of Intelligence Analysis* [26], a work on which we draw.

*CSLI, Stanford University, devlin@csli.stanford.edu. This research was supported in part by an award from the Advanced Research Projects Agency, under a subcontract to General Dynamics – Advanced Information Systems, as part of ARDA's NIMD Program.

1 Introduction

The importance of mathematics in military and defense-related matters has a long history, going back at least two-and-a-half thousand years: the famous ancient Greek mathematician Archimedes — who died at the hand of a Roman sword — devoted a considerable portion of his time and his extraordinary intellectual skill to military questions. Arguably the conflict with Germany in the Second World War was decided by the efforts of the American and British mathematicians, who were able to crack the German secret codes, and thereby maintain the crucial Atlantic shipping lanes that kept the Allies supplied. At the same time, the conflict with Japan was significantly shortened by the efforts of the many mathematicians and physicists involved in the Manhattan Project to develop the world's first nuclear weapon. Game Theory and Operations Research, two new branches of mathematics developed in the 1940s and 50s, played major roles (in addition to the continued importance of cryptography) in the Korean War, Vietnam, and the Cold War. More recently, the use of unmanned aircraft, guided missiles, and precision bombs in the two Iraq conflicts, in Bosnia, and in Afghanistan have also been heavily dependent on advanced mathematics.

The overall trend is clear. Over the centuries, and increasingly since around 1900, warfare and defense have made use not only of *more* mathematics but ever more *sophisticated* math, both in the design and deployment of weaponry and the never-ending struggle for superiority in intelligence gathering.¹

It is that latter category of intelligence gathering — or more precisely, intelligence analysis — that is the focus of this paper. For the United States, the events of September 11, 2001 marked the public start of another war,² the war on international terrorism. Unlike previous conflicts, where mathematics played crucial, but essentially *supporting* roles, the war on terror is primarily, indeed almost entirely, one of intelligence, with mathematics at the forefront. The only way to avoid more than a handful of major terrorist attacks on the United States and its citizens in the future — and anyone who claims we can completely eliminate the threat is surely either seriously deluded or else trying to be elected to public office — is through intelligence gathering and analysis. At present, several areas of mathematics play major roles in this effort, among them: Internet search techniques, data mining, decision theory, cryptography, automated text extraction, and statistical inference (particularly Bayesian inference, often implemented in the form of Bayesian networks). We can expect other, new branches of mathematics to develop in response to this new national need.

In addition to a new awareness of the more traditional elements of the intelligence gathering system, today's intelligence analyst has at his or her disposal a whole battery of sophisticated information systems, and new ones are being developed all the time. Indeed, the research reported in this paper was carried out as part of a team developing one such new support system. Those new systems are all heavily dependent on advanced mathematics. A natural question arises: does the utility of mathematics stop

¹It is often stated that the largest single employer of mathematics Ph.D.s in the world is the National Security Agency. Since the data is classified, this claim is impossible to verify, but it may well be true.

²It was already being waged out of the glare of the spotlights.

the moment the contractor ships the new information system — those sleek gray boxes and ever slimmer displays — to the government? Or can we bring some mathematics — or at least some mathematical insights — to bear on the next step of the intelligence analysis process, the part where the human analyst must weigh all the evidence and come up with a recommendation?

It would surely be naïve to imagine that mathematics could be brought to bear to such an extent that intelligence analysis reduced to solving an equation. Intelligence analysis is a complex human activity that, at least in my view, is intrinsically *not* reducible to an algorithmic (i.e., rule-based, step-by-step) procedure.³ But that does not mean mathematics cannot play any role; indeed, it can play a major one. It all depends on how the mathematics is applied.

There are (at least) two different ways that a mathematical analysis may be applied to a real world domain, in particular to an information processing domain. The more familiar way is that the mathematics forms the basis for a blueprint for a piece of engineering or a computer program. For example, a single mathematical algorithm developed by two Stanford graduate students in the 1990s was easily implemented (by them) as a computer program that formed the basis of a company the two students founded named “Google.”

The other way that mathematics may be applied is more indirect. A mathematical analysis of a particular domain provides a unique way of looking at the domain, a way that hides many features of the domain and emphasizes others. This results in a greatly simplified, yet logically sound view of the domain. (The hard part is doing it so that this view is also useful for the task in hand!) For example, aerospace physicists often view the solar system as a collection of point masses moving in a sea of gravitational forces. This is a highly simplistic and totally unrealistic picture of the true reality, but it’s just the right mathematically precise simplification you need to steer a spaceship to one or more of the planets. (This picture becomes unreliable when the spacecraft gets near to one particular planet. At that point another mathematical model must be used in its place.)

Given the nature of (the human parts of) intelligence analysis, with all its human complexities, it is the second, more indirect form by which mathematics is applied that we should expect in a study such as the one described here. In particular, when we start to throw mathematical symbols about, they are not intended to be read, absorbed, and understood by practicing intelligence analysts. Rather, they are part of our study of intelligence analysis. The symbolic descriptions we develop may be useful in themselves to the engineers who develop information support systems. But their influence on analysts, should such occur, will be through a heightened understanding of the process of analysis.

The intelligence community is already familiar with the utility of scientific study to improve its effectiveness through (in particular) Richards Heuer’s now classic work *Psychology of Intelligence Analysis* [26].⁴ What Heuer did was take scientific results

³See my book [14] for a detailed argument in favor of this claim.

⁴Re-published by the Central Intelligence Agency in 1999, currently available only on the Web.

discovered by the psychology research community and filter and interpret those findings for intelligence analysis. This paper has a broadly similar goal. Our principal aim is to investigate the degree to which mathematics may also be brought to bear on intelligence analysis.⁵ The particular focus of our research is the role played by context in intelligence analysis (and subsequent decision making).

In summary, a mathematical analysis of the way intelligence analysts work may be expected to lead to sharper insight into the human problems inherent in intelligence analysis, resulting in particular in:

- improved designs of information systems and other support tools;
- improved analysis protocols to guide the intelligence analysis community.

2 The 2002 decision to invade Iraq

It is now generally acknowledged that a key item of information behind the 2002 U.S. decision to invade Iraq was a single satellite photograph. With hindsight, it now appears that the interpretation of the photograph was incorrect. There can hardly be a more dramatic illustration of the crucial need for a scientific understanding of the way that people interpret (i.e., attach meaning to) symbols, be those symbols words, reports, diagrams, sketches, graphs, maps, PCR sequences (DNA prints) or, as in this case, photographs. (By ‘scientific’, we mean a mathematically precise understanding based on empirical study.)

On February 4, 2004, the newspaper *USA Today* had a major cover story⁶ headlined “A desert mirage: How U.S. misjudged Iraq’s arsenal.” According to the article, the key piece of evidence that tipped the scales in favor of a U.S. invasion of Iraq was the satellite photo shown in Figure 1.

As the *USA Today* article, together with several other examples we will present, illustrate, symbols (which, as above, may be words, reports, diagrams, etc.) on their own, do not have meaning. Rather, human beings (and perhaps, in some cases, other living organisms) *attach* meaning to symbols.⁷

According to the article:

The convoy photos . . . were decisive in a crucial shift by U.S. intelligence: from saying Iraq might have illegal weapons to saying that Iraq definitely had them.

⁵Of course, mathematics — particularly statistical analysis — plays an important role in most psychological studies, including the work described by Heuer. But we mean a more direct use of mathematics.

⁶The full story is reproduced in the Appendix.

⁷Expressing it this way gives the impression that symbols are somehow more basic than the interpretations we put on them. The relationship between symbols and their meanings is actually more of a symbiotic one. Something can only be a “symbol” if an agent attaches meaning to it. However, any symbol can — and does — have more than one meaning, and different symbols can have the same meaning. This entire paper amounts to a detailed analysis of this many–many symbiotic relationship.



Figure 1: The key satellite photo that led to the 2002 U.S. decision to invade Iraq. But what exactly does it show?

The assertion that Saddam had chemical and biological weapons – and the ability to use them against his neighbors and even the United States – was expressed in an Oct. 1, 2002, document called a National Intelligence Estimate. The estimate didn't trigger President Bush's determination to oust Saddam. But it weighed heavily on members of Congress as they decided to authorize force against Iraq, and it was central to Secretary of State Colin Powell's presentation to the United Nations Security Council a year ago this week.

...

Of all the Bush administration accusations about Iraq, none was more important than the charge that Saddam possessed chemical and biological weapons capable of killing millions of people. And no evidence was more important to making that charge than the convoy photographs taken in March, April and May 2002.

...

[the convoy images were] "an extraordinarily important piece. It's one of those 'dots' without which we could not have reached that judgment that Saddam had restarted chemical weapons production."

According to the *USA Today* article, "The story of the suspicious convoys in the Iraqi desert illustrates how the CIA turned tantalizing evidence of Iraqi weapons into conclusions that went beyond the available facts."

It is not the purpose of this research paper to evaluate each individual statement in the newspaper article. Rather we examine the role played by contextual factors in the evaluation of the satellite photograph as part of the decision making process that led to the U.S. invasion of Iraq, and use it, along with other examples, to motivate and inform a mathematical analysis of the role of context in intelligence analysis. The

mathematical analysis we present will stand regardless of the accuracy of any specific claim in the *USA Today* report.

By way of background, the story in *USA Today* led off as follows.

One year before President Bush ordered the invasion of Iraq, a U.S. spy satellite over the western Iraqi desert photographed trailer trucks lined up beside a military bunker. Canvas shrouded the trucks' cargo.

Through a system of relays, the satellite beamed digitized images to Fort Belvoir in Virginia, south of Washington. Within hours, analysts a few miles away at CIA headquarters had the pictures on high-definition computer screens. The photos would play a critical role in an assessment that now appears to have been wrong – that Iraq had stockpiled weapons of mass destruction.

The way analysts interpreted the truck convoy photographed on March 17, 2002 – and seven others like it spotted over the next two months is perhaps the single most important example of how U.S. intelligence went astray in its assessment of Saddam Hussein's arsenal. Analysts made logical interpretations of the evidence but based their conclusions more on supposition than fact.

The eight convoys stood out from normal Iraqi military movements. They appeared to have extra security provided by Saddam's most trusted officers, and they were accompanied by what analysts identified as tankers for decontaminating people and equipment exposed to chemical agents.

The article lists a number of contextual factors that were instrumental in the way the intelligence service ultimately interpreted the photograph. Perhaps most significant of all was the fear of a repeat of the failure to predict the terrorist attack of September 11, 2001:

- Despite the lack of proof, CIA Director George Tenet and his top advisers decided to reach a definitive finding. Based on experience with Iraq and with the Sept. 11, 2001, terrorist attacks in mind they were far more worried about underestimating the Iraqi threat than overestimating it.

According to the article, “[The key decision makers] were haunted by past failures and the fear of the worst-case scenario.”

Other contextual factors that came into play were:

- Scarcity of evidence [to have gotten rid of WMD] stemmed not from innocence but from Iraqi concealment and lies.
- ... any finding exonerating Iraq would put them into conflict with top administration officials.
- U.S. intelligence analysts were reluctant to give Iraq the benefit of the doubt because Saddam had fooled them before. After the 1991 war, U.N. weapons inspectors, tipped off by an Iraqi defector, uncovered a much more extensive program to develop nuclear weapons than the CIA had estimated. It happened again in 1995 when Iraq admitted to a biological weapons program undetected by U.S. intelligence.
- Virtually all of the CIA's recent, painful lessons revolved around the failure to detect and warn of a threat. These included a bombing at the Khobar Towers military barracks in

Saudi Arabia in 1996; nuclear tests by India and Pakistan in 1998; the bombing of the USS Cole in Yemen in 2000; and, most traumatically, the Sept. 11 attacks.

- Once Iraq showed it knew how to make chemical weapons in the 1980s, U.S. intelligence assumed it held on to the recipe. “Iraq’s knowledge base is absolutely critical,” Cohen said. “Knowledge is not something you lose.”

One feature of the intelligence gathering that the news article implies is that the intelligence services engaged in considerable conscious, deliberative reasoning to arrive at their interpretation of the satellite photograph — possibly over several days or weeks. This is quite different from the instantaneous interpretation of a symbol that takes place, for example, when an experienced car driver comes up to a red light, and knows immediately that it means STOP. In our recent paper Devlin [16], we analyzed such evidence-based, context-influenced reasoning. According to the report in *USA Today*, some contextual features actively and directly influenced the process of reasoning itself:

- In July 1998, a commission led by Donald Rumsfeld had advised adoption of a new kind of analysis to “extrapolate a program’s scope, scale, pace and direction beyond what the hard evidence at hand unequivocally supports.”
- “We put the analysts under tremendous pressure,” said Kay, the former head of the post-war weapons search. “There is a point where an analyst simply needs to tell people: ‘I can’t draw a conclusion. I don’t have enough data. Go get me more data.’ But in the wake of 9/11, believe me, that is difficult to do.”

3 The six Ps of context

With the Iraq invasion decision to hand now, let’s take an initial look at the way context influences the way we interpret, say, a photograph (and hence affects analysis and decision making).

The first thing to note, and this is crucial, is that the process by which an agent attaches meaning to a symbol *always* takes place in a context, indeed generally several contexts, and is *always* dependent on those contexts.

An analytic study of the way that people interpret symbols comes down to an investigation of the mechanism captured by the diagram:

$$[\text{AGENT}] + [\text{SYMBOL}] + [\text{CONTEXT}] + \dots + [\text{CONTEXT}] \longrightarrow [\text{INTERPRETATION}]$$

Much of this paper will be an analysis of just what a context is and how it affects cognitive activity. By way of setting the scene, let’s note that contexts do not exist in isolation; rather, something is a context *for some particular action*.

As a preliminary definition, we may say that a feature \mathcal{F} is *contextual* for an action \mathcal{A} if \mathcal{F} constrains \mathcal{A} and may affect the outcome of \mathcal{A} , but is not a constituent of \mathcal{A} . In particular, a situation, environment, or set of circumstances C is said to be a *context* for an action \mathcal{A} if C constrains \mathcal{A} or affects the outcome of \mathcal{A} but is not a constituent of \mathcal{A} .

Five observations about contexts motivate and guide the research presented in this paper. Those observations can be summed up by what we call *the six Ps of context*:⁸

1. Contexts are **p**ervasive
2. Contexts are **p**rimary
3. Contexts **p**erpetuate
4. Contexts **p**roliferate
5. Contexts are **p**otentially **p**ernicious

First of all, contexts are pervasive. Every object that exists is situated in a context and everything that happens does so in a context. In fact, everything that exists does so in multiple contexts, and every action takes place in multiple contexts.

The second P, that contexts are primary, takes note of the fact that when any action takes place, it does so in a context that exists prior to that action. In fact, in general actions take place in multiple contexts, each of which exists prior to the action. Some contexts for an action may exert a significant influence on the action, and in general those are the contexts that we are likely to identify *as contexts* for the action.

In many situations, actions occur in a sequence, with each one dependent on those that went earlier. This is true, in particular, of much of intelligence analysis, where each step in the analysis provides part of the basis for subsequent steps. In such situations, the completion of one action \mathcal{A} in a context \mathcal{C} establishes a new context, which we may denote by $\mathcal{C}[\mathcal{A}]$, for the next action. In this way, contexts perpetuate. This feature has the potential to result in either very good conclusions or very bad ones.

A naïve approach to trying to handle context is to identify the key contextual features, make them precise, and, in the case of a formal analysis, formalize them. But this fails to properly appreciate what contexts are and how they affect actions, and is doomed to fail.

For instance, as Garfinkel showed in [20] — and as we emphasized in [14] (which presented examples from [20]) — the process of trying to make context precise is open ended. Like trying to kill Hydra by cutting off one head, pulling one or two features from the context and handling them formally leaves you facing a whole new range of contexts for the features pulled out. This point is so important, and so often ignored, that we present here one of the many examples Garfinkel presented to support his observation.

In this example, Garfinkel asked the students in one of his classes at UCLA to write a report on an everyday conversation. The next time they found themselves in a conversation, he instructed them, they were to note down everything that was said. When the conversation was over, they were to add explanations to their notes, describing the contextual information behind their statements. The resulting report

⁸One of the slogans has two P's; hence five slogans.

would then provide an account of what was actually said, what was assumed (the context), and what was meant.

When the class next met, one of the students reported on the following conversation between himself and his wife. What was actually said is written in Roman type; the speaker's subsequent interpretation follows in italics.

HUSBAND: Dana succeeded in putting a penny in a parking meter today without being picked up.

This afternoon as I was bringing Dana, our four-year-old son, home from the nursery school, he succeeded in reaching high enough to put a penny in a parking meter when we parked in a meter zone, whereas before he has always had to be picked up to reach that high.

WIFE: Did you take him to the record store?

Since he put a penny in a meter that means that you stopped while he was with you. I know that you stopped at the record store either on the way to get him or on the way back. Was it on the way back, so that he was with you or did you stop there on the way to get him and somewhere else on the way back?

HUSBAND: No, to the shoe repair shop.

No, I stopped at the record store on the way to get him and stopped at the shoe repair shop on the way home when he was with me.

WIFE: What for?

I know of one reason why you might have stopped at the shoe repair shop. Why did you in fact?

HUSBAND: I got some new shoe laces for my shoes.

As you will remember I broke a shoe lace on one of my brown Oxfords the other day so I stopped to get some new laces.

WIFE: Your loafers need new heels badly.

Something else you could have gotten that I was thinking of. You could have taken in your black loafers which need heels badly. You'd better get them taken care of pretty soon.

A number of things are obvious about this particular exercise. First, the original conversation is remarkably everyday and mundane, and concerns an extremely restricted domain of family activity. Second, the degree of detail given in the subsequent 'explanations' or 'elaborations' of what each person said seems quite arbitrary. It is easy to imagine repeating the exercise over again, this time providing still further explanation. And then it could be repeated a third time. Then a fourth. And so on, and so on, and so on. Apart from boredom or frustration, there does not seem to be any obvious stopping point.

Indeed, this issue of how much detail to provide was raised by the students themselves. As Garfinkel himself reported [20, p.26]:

Many students asked how much I wanted them to write. As I progressively imposed accuracy, clarity, and distinctness, the task became increasingly laborious. Finally, when I required that they assume I would know what they had actually talked about only from reading literally what they wrote literally, they gave up with the complaint that the task was impossible.

The dilemma faced by Garfinkel’s students was not simply that they were being asked to write ‘everything’ that was said, where that ‘everything’ consisted of some bounded, albeit vast, contextual content. It was rather that the task of enumerating what was talked about itself extended what was talked about—the horizon of understanding continued to recede with every cycle of increased explanation. Quite simply, the task was endless. At every stage, what has been specified is dependent on further contextual factors.

As Garfinkel’s experiment makes clear, every action is in at least one context that cannot be ignored. Consequently, the solution to the context problem is *not* simply to identify key features of the context and make them part of the formal reasoning. This is not to say that this should not be done. Indeed, sometimes this is a key part of the solution. But note the appearance of that key term ‘part’ in that last sentence. Bringing a key feature of the context into the formal system is what Garfinkel’s students did in each iteration of their exercise.

The point to recognize is that contexts are inescapably *outside the system* (whatever you take as the system). (That last statement is actually a tautology, but one that we feel is worth making.) Hence, reasoning that takes account of context is inescapably *two-sorted*. The most that can be done is to recognize, flag, and track key contexts; they cannot be eliminated by “pulling them into the formal reasoning.” Pulling in features of contexts simply leads to rapid context proliferation, a P-factor that is already a challenge without adding further to it in a misguided attempt to sidestep the issue.

This last observation will have profound, limiting impact on the way we apply mathematics to the problem and the kind of mathematics we can utilize.

Finally, as we shall see in the examples that follow, when contexts proliferate, their influences can be potentially highly pernicious — the final two Ps.

Referring back to the *USA Today* article on the Iraq invasion decision, notice that, although each of the features we identified as contextual did influence the process of interpreting the photograph, none of them were actually part of the act of interpretation. They were *contextual*. It could have been that, at some stage of the analysis, one of the analysts said “Look guys, Secretary Rumsfeld has advised us to go beyond what the hard evidence at hand unequivocally supports.” Such a statement would have brought that feature into the reasoning process itself, but it would still have remained (in addition) a feature of the context in which the reasoning took place. In our paper Devlin [16] we provide mathematical machinery for tracking such dual uses of features of context. This machinery is two-sorted, embodying both a formal reasoning component and a mechanism to flag, tag, and track contexts. (Actually, our entire framework is formal. The distinction we are making here is between the entities the intelligence analyst treats as formal objects and the contexts both of those entities and

within which the analyst operates.)

Another example of a deliberative process of evidence-based, context-influenced reasoning, that has many similarities with the Iraq invasion decision, is provided by NASA's decision — in the event disastrous and hence subsequently controversial — to launch the Space Shuttle Challenger in January 1986, analyzed at length by Diane Vaughan in her book [41]. In that case, arguably the most crucial of several contextual factors was intense pressure on NASA from the White House to “put on a good show.” (The flight was a much publicized one in which an “ordinary schoolteacher” was one of the crew; her inclusion was intended to provide the nation's schoolchildren with a great role model.)

Both the Iraq and the Challenger launch decisions provide dramatic evidence for the need for a study such as the one we are embarking on here. From the point of view of developing a mathematical framework for analyzing such decision making, however, both examples turn out to have too many complexities to be taken as starting points. A more tractable example of the role played by context in the assignment of meaning to symbols — and the problems that can arise — is provided by an earlier decision to launch a military action: the nineteenth century Charge of the Light Brigade, immortalized by Alfred Tennyson in his epic poem of that title. That disastrous military action arose because of the contextual misinterpretation not of a photo but of a single word.

4 The Charge of the Light Brigade

In March, 1854, following Russia's attempts to expand her empire, Britain and France had declared war on Russia. On 14 September, British and French forces invaded the Crimea, a northern peninsula on the Black Sea that is today a part of the Ukraine. The aim was to capture Sebastopol and destroy the Russian fleet. Following a major victory at the Alma, when a 40,000 strong Russian force was defeated, the allies advanced toward Sebastopol, putting it under siege on 17 October. The British troops were led by Fitzroy Somerset, Lord Raglan.

On 25 October, the Russians attempted to break the siege, launching an attack on the British positions near Balaclava, a small town 6 miles to the south-east of the city. The attack was repelled by the guns of the British Heavy Brigade. Raglan ordered the cavalry, his Light Brigade, to move in and sweep the enemy from their redoubts while they were still in disarray. His order to Lord Lucan, the commander of the Light Brigade, read:

Cavalry to advance and take advantage of any opportunity to recover the Heights.
They will be supported by the infantry, which have been ordered to advance on two fronts.

When Lucan received the order, he took it to mean that he should begin his advance only when the supporting infantry arrived. Seeing none, he waited. Meanwhile, the

Russians began to recover their positions, and started to drag away captured British guns.

Forty-five minutes later, Raglan — who from his hilltop position could see what the Russians were doing — asked his aid Airey to send a second message. It read:

Lord Raglan wishes the cavalry to advance rapidly to the front follow the enemy and try to prevent the enemy carrying away the guns. Troop Horse artillery may accompany. French cavalry is on your left. Immediate. Airey.

Airey gave the note to his ADC, Captain Nolan, to deliver. As Nolan sped away on horseback, Raglan shouted after him, “Tell Lord Lucan the cavalry is to attack immediately.”

Lucan read Airey’s message. It made no sense. From his position down in the valley, he could see neither the enemy nor any guns, apart from the massive Russian force holding what looked like an invincible position at the far end of the valley. Seeing Lucan’s indecision, Nolan passed on Raglan’s verbal message: “The cavalry is to attack immediately.”

The order seemed clear. Under Lucan’s command, the 673 soldiers of the Light Brigade started to make their way down the valley toward the Russian position. In the carnage that followed, 272 of them were killed. They never had a chance. What is more, they were never supposed to advance on such an impregnable position. From Raglan’s elevated position on high ground, he looked down on the routed Russians, pulling back and taking British guns with them. That was the “front” he meant when he issued the order for the cavalry to “advance rapidly to the front.” But the only enemy Lucan could see was the one at the end of the valley. That was his “front.” And when, like a good soldier, he finally agreed to obey the order he had been given, that was the front he attacked. The valley along which he advanced would become Tennyson’s “Valley of Death.”

Their advance stalled, the British troops had to spend the entire winter holed up in the Crimea, and it was not until September of the next year that they finally took Sebastopol. The war ended the following February. Altogether a considerable cost to pay for misunderstanding the referent of the single word “front.”

One word, which everyone understood correctly; two different interpretations, depending on different contexts. If Lucan and Airey had been able to engage in a normal conversation, almost certainly the misunderstanding and the ensuing tragedy would have been avoided. But given the actual circumstances, with Lucan and Airey in different locations and Raglan and Nolan acting as intermediaries, to say nothing of everyone being in the heat of battle, it is not at all hard to understand how things went wrong.

We can represent the problem diagrammatically as in Figure 2. In such diagrams, we represent symbols by enclosing them in rectangles and show contexts for the interpretation of those symbols as ovals that enclose those rectangles. In this case, the single symbol FRONT is enclosed in two ovals, indicating the two different contexts in which

the word was interpreted, leading to the two different real world meanings (*referents*) attached to the word.

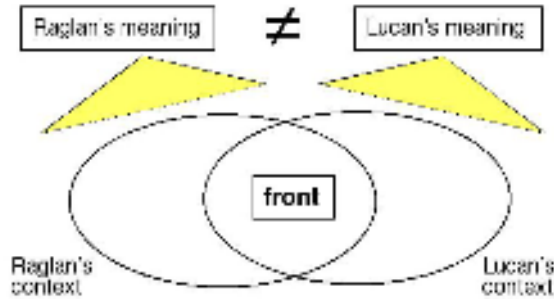


Figure 2: The two interpretations of the word “front” that led to the disastrous charge of the Light Brigade.

Before we go any further, it should be noted that misinterpretation of a particular word or phrase is not an unusual occurrence. Rather, it’s the norm. All words significantly underspecify their interpretation. This gives rise to what Barwise and Perry [4] called *the efficiency of language*, whereby a relatively small number of words is sufficient to refer to and talk about an unlimited number of objects, people, other living entities, and events. Words constrain their possible interpretations, but they do not specify them. It is (only) when a word is used, in a specific context, that it acquires its interpretation.

In face-to-face conversation, which is the environment in which language evolved and for which it is (therefore) best suited, we have developed a great many linguistic and extra-linguistic signals and other mechanisms to ensure that, for the most part, the participants agree on the interpretations each puts upon the words exchanged. Our Stanford colleague Herb Clark calls this *maintaining the common ground* of a conversation. (See, for example, his book Clark [11].) These mechanisms work well in a face-to-face setting, less well — but still tolerably well — when two people speak on the telephone, and considerably less well when three or more people take part in a telephone conference call. When the conversation takes place through an intermediary — as occurred with Raglan and Lucan’s exchange just prior to the charge of the Light Brigade — or else in the nowadays common situation of computer mediated communication, things can go badly wrong, sometimes disastrously so. The common ground can easily be lost, causing divergent interpretations to emerge.

The only possibility for resolving the problem of maintaining common ground, or at least minimizing and mitigating breakdowns of the kind illustrated by this last example, is to design our systems to incorporate as much as possible the feedback failsafe mechanisms of human–human communication and human societies. With human–machine interaction or mediated human–human communication, this is not easy. Because the communication between Raglan and Lucan was mediated by the courier Airey, the two commanders were unable to recognize and rectify the fact that they had formed differ-

ent interpretations of the word “front”, even though Lucan tried not once but *twice* to implement corrective feedback.

Couple context bifurcation with the perpetuation effect, where each stage in the communication provides a context for all the following stages, and very soon a branching tree of different contexts can develop, with the participants in the communication each following their own path through that tree. This is how the fourth P-factor can develop: context proliferation.

And it isn’t just people who make decisions in contexts. Computer systems — indeed technologies in general — also operate in contexts. Sometimes the consequences of technological context differences can be just as deadly as in the human case, as our next example shows.

5 The crash of flight AA 965

In December 1995, American Airlines flight 965 from Miami to Colombia was on its final approach to Cali airport, when it went off course and crashed into a nearby mountain range, killing all 159 passengers and crew on board.

When the report of the airline’s investigation was made public in August of the following year, it became clear that the problem was not mechanical. Nor was the weather a factor: there was a lot of cloud, but with modern navigational systems that is not a problem. The principal culprit was information; more precisely, a contextual effect on the data provided by the on-board computer system that led the pilot to form a different interpretation to the aircraft’s onboard guidance system.

Here is how a senior executive of the airline (its “chief pilot”) subsequently described the events that led up to the crash. The air traffic controller at Cali instructed the crew to fly toward a nearby beacon called “Rozo”, identified on navigational charts by the letter R. The crew entered that letter into the on-board flight management computer, whereupon the screen responded with a list of six navigational beacons. By convention, such a list normally presents the beacons ranked from nearest to farthest from the plane. Since the crew was on the final approach, they did the customary thing and accepted the top entry on the list. It should have been the Rozo beacon, in agreement with the convention used on the printed charts. It was not.

Unknown to the crew, the R at the top of the list actually signified a beacon called “Romeo” in Bogota, more than 100 miles away and in a direction more than 90 degrees off course. Once the crew had selected the R-beacon on the computer, the autopilot silently and obediently did as instructed, and slowly turned the plane left toward Bogota. By the time the crew realized something had gone very wrong, it was too late.

Officially, there was no question as to who was at fault. It was the crew’s job to know where the plane was headed and what the autopilot was doing. But when alert, well-trained personnel can make such a catastrophic error, we should try to see what circumstances led them to do so.

In the case of flight AA 965, one obvious and vital factor is the importance of consistent and accurate data. On the charts, the Rozo beacon was labeled “R”. But in order to retrieve its listing from the computer, the crew would have had to type the entire word “Rozo”. The airline’s accident report did not explain the discrepancy, but the chief pilot’s report noted “charting and database anomalies that have been discovered.”

A more general issue is how to present important information to those who need it. Given the myriad duties of an airline crew as they come in to land, they have little if any opportunity to double check every single detail of the masses of information available to them. Aware of the dangers of providing busy crews with too much information, cockpit displays are designed to supply only the really essential information, and moreover to do so in as simple and concise a manner as possible. Likewise, it makes sense to arrange things so that the crew do not have to enter a complete word if one or two letters is sufficient.

As an American Airlines spokesman explained, the screens in the cockpit show only the beacons’ code letters and geographical coordinates. Since the corresponding charts generally show those coordinates in print so tiny a busy crew is unlikely to check them indeed, they may even omit them entirely the crew will almost certainly rely on the letter or name used to identify the beacon.

Figure 3 indicates the source of the problem in terms of contexts. The crew were operating within two different sets of conventions. The problem was, they were not aware of the fact. They thought they were operating under the normal conventions for abbreviating the names of landing beacons. According to the conventions in that context, the letter R denoted the Rozo beacon. Their assumption was that the local computer system used the same convention. If that were the case, then the letter R would denote the Rozo beacon on the computer as well. However, the people who programmed the local computer system had a different context, in which the letter R denoted the Romeo beacon.

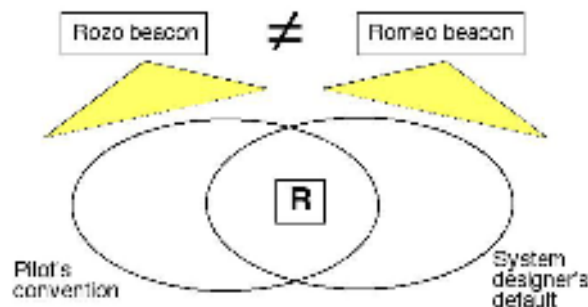


Figure 3: The two interpretations of the letter R that led to the fatal crash of flight AA 965.

Of course with hindsight, it is easy to see how to minimize the likelihood of a repeat of the flight 965 accident: either make sure that the abbreviations used by

the computer system are the same as those on the printed charts, or else arrange for the cockpit display to show the complete beacon name (“Rozo” or “Romeo” in the case in question) along with the coordinates. Drawing a diagram such as Figure 3 does not, in itself, solve the problem. What the diagram does do is highlight exactly where the problem lies, namely the context. After all, there was only one crew, only one computer system, and only a single keystroke at the crucial moment, the letter R. The two contexts were largely in agreement (indicated by their being depicted as overlapping ovals). But one crucial difference between them was the actual beacon associated with the letter R. The convention that made this association should have been in the overlapping region, but it was not.

In our book *Infosense* [15], from which the two examples above are taken, we present several more examples to show how context affects the interpretation of symbols. Although this paper is self-contained, reading it will probably proceed much more quickly from now on if that book is read first.

Finally, by way of an illustration of the degree to which contexts are *pervasive*, note that even numbers are susceptible to context effects. In 1999, NASA lost a \$125 million Mars orbiter because a Lockheed Martin engineering team used English units of measurement while the agency’s team used the more conventional metric system for a key spacecraft operation.

Lockheed Martin helped build, develop and operate the spacecraft for NASA. Its engineers provided navigation commands for Climate Orbiter’s thrusters in English units, but NASA used the metric system. The units mismatch prevented navigation information from transferring between the Mars Climate Orbiter spacecraft team at Lockheed Martin in Denver and the flight team at NASA’s Jet Propulsion Laboratory in Pasadena, California.

The numbers themselves were accurate. Each piece of software in the guidance and control system operated exactly as intended. The problem was not on the spacecraft. It was entirely a problem of context, in the form of two different contexts that were established long before the mission began. Those two different contexts were operating — on the ground — throughout the entire orbiter development and construction process. But on the ground, this was not a problem. It was in flight that the problem arose. From the moment it was in flight, the orbiter was operating in both contexts, one where numbers are interpreted according to English units, the other where numbers are interpreted according to the metric system. As the orbiter approached Mars, the influences of those two contexts came into conflict, with disastrous consequences for the mission.

Before we go on, we should stress that the four examples we have presented, the Iraq invasion decision, the charge of the Light Brigade, the crash of AA 865, and the loss of the Mars orbiter, are exceptional only in their catastrophic outcomes. The context issues that led to those dramatic outcomes are not only familiar but the norm. Most of the time, evolved or designed feedback mechanisms and failsafe procedures keep things running fairly smoothly and paths to disaster are cut off at an early stage.

6 Intelligence analysis: two views from the air

Having recognized the role played by context in intelligence analysis, and other activities, and made some observations of the manner in which contexts can condition and influence actions, the next step is to try to develop some mathematical apparatus to begin a more systematic analysis.

First, we need to clarify exactly what we mean by a *mathematical* analysis. Mathematics gains its enormous analytic and predictive strength from three features, that can give rise to three different kinds of analysis.

- The use of numbers and statistical techniques can give rise to analyses of aggregate behaviors exhibited by sufficiently large populations.
- The use of formulas, equations and numbers can give rise to enormous precision of description.
- Abstraction can lead to deep insights and a recognition of similarities with other domains.

The approach we adopt in this paper is very much of the third kind, abstraction. It is highly likely that, in due course, the techniques developed here, when they reach a suitable stage of maturity, may be combined with statistical methods to provide more quantitative results than the essentially qualitative ones we shall obtain. But we suspect that, in general, the nature of individual or small group human activities, such as intelligence analysis or the subjects they study, are not amenable to the kinds of equation-rich, precise mathematical analyses found in, say, physics. In other words, what we are doing is perhaps more accurately described as adopting a *mathematical approach* to a study of intelligence analysis, based on abstraction, rather than *applying mathematics* to that domain. We have elsewhere referred to this kind of approach as “soft mathematics.” (See Devlin [14].)

One way to think of the process of abstraction is in terms of flying over the domain in an aircraft and viewing the activity from greater and greater altitudes. As you go higher, details become lost, allowing the truly significant features to stand out more. Moreover, the higher you go (the greater the degree of abstraction), the further you can see, sometimes enabling you to recognize important similarities between two regions that you previously thought were quite different.

Of course, this process is not always guaranteed to lead to useful results. If the answer can be found *only* in the details, it's best to remain on the ground. But in most cases, a view from above can be useful, even if only to augment information gleaned on the ground. As has been amply demonstrated on several occasions, provided your targeting is sufficiently accurate, even from 35,000ft you can inflict an awful lot of damage on the ground.

With this perspective in mind, Figure 4 shows a schematic overhead view of an intelligence analyst from an altitude of 20ft. The principal feature of the analyst's



Figure 4: The intelligence analyst at work: the view from 20ft.

activity that the diagram shows is that he⁹ acquires information from several different sources, which present that information in several different modalities.

Reasoning with information presented by different media is known as heterogeneous reasoning, and has been studied extensively, both at CSLI (where I work) and elsewhere. See, for example, Barwise and Etchemendy [5], [6], [7], [8] and Devlin [16]. In particular, [16] provides a framework for analyzing reasoning in such an environment.

Figure 5 shows a schematic overhead view of an intelligence analyst from an altitude of 20,000ft. This represents the configuration that motivated the analysis we presented in [16]. The two ovals labeled “source” are intended to represent human sources or human organizational sources (as opposed to the more amorphous, massive data sources denoted by clouds).

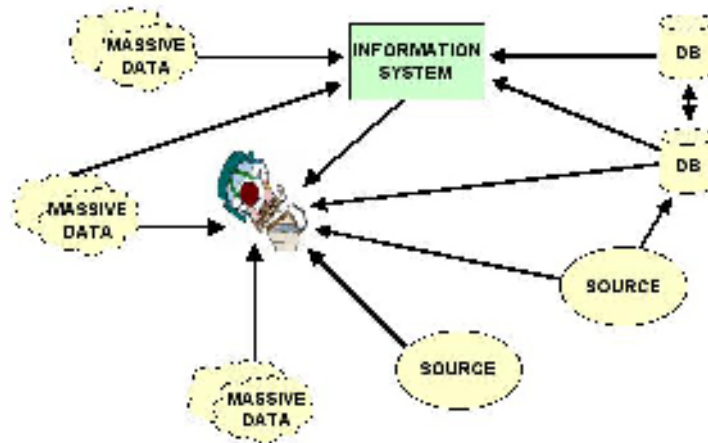


Figure 5: The intelligence analyst at work: the view from 20,000ft.

⁹This analyst is male by way of the clipart provided by Microsoft.

7 A logic of context-influenced reasoning

The principal goal of the research presented in [16] is to inform the development of the automated information system indicated by the rectangular box in Figure 5 (including the interface presented to the analyst through his desktop workstation). But note that our paper analyzed the *entire system* depicted in Figure 5, including the human analyst. This is quite different from traditional attempts to apply mathematical techniques to this domain. Typically, the human is viewed as a *user* of an inanimate system. Mathematics is used to describe the flow of data around the computer network, cognitive psychology is used to study the activity of the human user, and HCI bridges the gap between the two subdomains. Our mathematics applies to the whole system, human(s), machines, and network.

In connection with that last remark, for readers familiar with work done on formalizing contexts within the artificial intelligence community, we should point out that the research presented here has a very different goal. The framework we develop below certainly has explicit representations of context. Nevertheless, it is not an attempt to develop another “logic of context,” at least as that term is commonly understood. Trying to handle context in formal logic was first introduced as a research thread in artificial intelligence by John McCarthy in his 1971 Turing Award lecture, subsequently published as [29]. See also McCarthy’s more recent article on the subject [30]. Many other attempts have been made to incorporate features of context in formal logic, for example Attardi and Simi [1], Buvac and Mason [10], Farquhar and Buvac [18], Farquhar, Dappert, Fikes, and Pratt [19], Giunchiglia [22], Guha [23], and Shoham [36]. The microtheories employed by the Cyc system are also clearly an attempt to take account of context. In addition, John Sowa summarizes several attempts to formalize context in logic and in natural language semantics in his online paper [37], and Rich Thomason [38] is engaged in an ongoing project to formalize context.

Our approach is quite different. While we recognize the utility of computer reasoning systems that can reason in specific, defined “contexts” — we would prefer to call them *local domains* — for reasons we have already articulated, we firmly believe that the influence of context cannot be fully accounted for this way. Rather, we start from the position that, when a trained human analyst, working with one or more computer information systems, reasons about real world affairs in a domain in which she or he is expert, the human analyst is in general far better able to make key decisions about matters of context than the automated system. Thus, rather than reify contexts and integrate them into a formal system, as previous work on context logics has done, we set out to build a framework that has two distinct parts, a formal system and a context-tracker. The contexts that the context-tracker represents are not reifications as mathematical objects but pointers to real-world contexts, with all that entails. Our framework sets out to capture, as far as possible, the linkages between, and influences of, context and the elements of the formal system. As a model of an activity, our “logic” attempts to model *the human reasoner* — perhaps aided by one or more information systems, possibly including automatic reasoners. It is not intended to form a blueprint for the design of an automated system. It may, of course, *inform* the design of such a

system, and indeed that is one of our principal goals. In which case, we would expect that system to make use of some “context logic,” such as one of those developed by the researchers referred to above. Our framework may also form a basis for various reasoning and reporting protocols to be followed by analysts.

In the framework we developed in [16], the degree of confidence the analyst places upon each piece of information being considered is tracked by informational flags (which may be numeric or qualitative), which the agent considers during each stage of the reasoning. We view evidence-based reasoning as a temporal cognitive process that acts not on statements σ (as in the standard mathematical model of reasoning) but on entities of the form

$$s \models_{\tau_1, \tau_2, \dots} \sigma$$

where:

1. σ is a statement (or fact);
2. s is a *situation* (in the sense developed at CSLI during the period 1983–93 and described in Devlin [13]) which provides support or context of origin for σ ; and
3. τ_1, τ_2, \dots are the *indicators*¹⁰ of σ , i.e., the specific items of information in s that the reasoner takes as justification of σ .

We call an entity of the form $s \models_{\tau_1, \tau_2, \dots} \sigma$ a *basic reasoning element*.

Within that framework, a process of evidence-based reasoning can be represented like this:

$$\begin{array}{ccc} s_1 & \models_{\tau_1, \dots} & \sigma_1 \\ s_2 & \models_{\tau_2, \dots} & \sigma_2 \\ s_3 & \models_{\tau_3, \dots} & \sigma_3 \\ & \vdots & \\ \hline s & \models_{\tau_1, \dots, \tau_2, \dots, \tau_3, \dots} & \sigma \end{array}$$

where each basic reasoning element either supplies evidence for the reasoning or else follows from one or more previous elements by a logical deduction rule.

More formally, by an *evidential reasoning element* we mean a 1×3 matrix of the form

$$\boxed{\text{FACT} \mid \text{SUPPORT} \mid \text{INDIC}(1), \text{INDIC}(2), \dots}$$

such that

$$\text{SUPPORT} \models_{\text{INDIC}(1), \text{INDIC}(2), \dots} \text{FACT}$$

By an *evidential reasoning step* we mean a finitary array of the form

¹⁰Our use of the term “indicators” with this meaning comes from social science.

| | | | |
|----------|---------------------|------------------------|---|
| OPERATOR | FACT ₁ | SUPPORT ₁ | INDIC ₁ (1), INDIC ₁ (2), ... |
| | FACT ₂ | SUPPORT ₂ | INDIC ₂ (1), INDIC ₂ (2), ... |
| | ... | | |
| | FACT _k | SUPPORT _k | INDIC _k (1), INDIC _k (2), ... |
| OUTPUT | FACT _{k+1} | SUPPORT _{k+1} | INDIC _{k+1} (1), INDIC _{k+1} (2), ... |

where each row

| | | |
|-------------------|----------------------|---|
| FACT _i | SUPPORT _i | INDIC _i (1), INDIC _i (2), ... |
|-------------------|----------------------|---|

is an evidential reasoning element. The index k depends on the operator OPERATOR, and is called the *arity* of the operator.

The idea is that a basic evidential reasoning step consists of the application of the logical operator to one or more constituents of the evidential reasoning elements in its *scope* (the first k elements listed) to produce the *output element* in the final row.

An *evidential reasoning process* is a finite sequence ρ_1, \dots, ρ_n of basic reasoning steps such that each element is either *evidential* (i.e., an input to the reasoning process) or else the output of some previous (in the sequence) evidential reasoning step, or else is the special element STOP, which is the final element in the process. (STOP is a failure condition.)

The sequence of elements in an evidential reasoning process are not intended to provide a temporal model of the actual steps carried out by a reasoner. Rather, an evidential reasoning process models the logical flow of the reasoning as it leads to the conclusion. Much real-life reasoning is not linear. However, the model developed in [16] is such that any linear progression of steps in the actual reasoning a human carries out will be mapped to a linear ordering of the corresponding basic reasoning elements in the model.

Some of the actual operators that arise in any particular instance of evidence-based reasoning will depend on the specific circumstances that pertain in that application. We illustrate the general form the rules take by giving two of the more general ones.

Evidential Conjunction Rule

| | | | |
|---------|------------------------|----------------------------|--|
| CONJOIN | σ | s | τ_1, τ_2, \dots |
| | θ | t | $\gamma_1, \gamma_2, \dots$ |
| OUTPUT | $\sigma \wedge \theta$ | $s \cup t \cup \{\delta\}$ | $\delta, \tau_1, \tau_2, \dots, \gamma_1, \gamma_2, \dots$ |

where $\delta = \text{Con}\{\tau_1, \tau_2, \dots, \gamma_1, \gamma_2, \dots\}$, the assertion that the set $\{\tau_1, \tau_2, \dots, \gamma_1, \gamma_2, \dots\}$ is logically consistent (i.e., has no internal contradictions), and where the rule may be applied only if δ is valid. The restriction that δ is called the *indicator consistency condition* for the rule. If this condition is not satisfied, the rule produces the output STOP.

Evidential Modus Ponens Rule

| | | | |
|--------|-----------------------------|----------------------------|--|
| MP | σ | s | τ_1, τ_2, \dots |
| | $\sigma \rightarrow \theta$ | t | $\gamma_1, \gamma_2, \dots$ |
| OUTPUT | θ | $s \cup t \cup \{\delta\}$ | $\delta, \tau_1, \tau_2, \dots, \gamma_1, \gamma_2, \dots$ |

where $\delta = \text{Con}\{\tau_1, \tau_2, \dots, \gamma_1, \gamma_2, \dots\}$, and where the rule may be applied only if δ . If this condition is not satisfied, the rule produces the output STOP.

We need to exercise care in using these two rules. If the supports s and t are identical, there is in general no problem, nor if one support is contained within the other. In either of these cases, the indicator consistency condition can generally be assumed to be automatically satisfied, since reasoning generally proceeds under the tacit assumption that each individual source is internally consistent. (If, however, the reasoner suspects — or comes to suspect — that one of the supports used in the reasoning is internally inconsistent, then resolving that inconsistency becomes part of the reasoning process.)

The idea behind the approach in [16] is this. Coupling a fact σ with its support s in our framework does two things: (i) it acknowledges that σ does come from a particular source, and (ii) it provides a record of that source. Explicitly listing the indicators τ_1, τ_2, \dots with σ and s puts on record the particular items of information in s that the reasoner believes are salient in supporting σ , and uses to justify making use of σ in the reasoning. When an unexpected or troublesome conclusion is reached, or when the reasoning fails to yield a conclusion, it may be necessary to re-examine the veracity of some of the facts used in the reasoning, and that may involve reconsideration of the indicator already identified, or a search for indicators hitherto ignored. In an extreme case, the reasoner may have to question an entire source, perhaps rejecting it and looking for evidence elsewhere.

For further details we refer to [16].

8 Are we looking in the right place?

The disaster examples we presented in the early part of the paper illustrated that:

1. context is always present
2. context effects are frequent and often decisive
3. even when you identify a particular context, it is generally plastic and ineffable and impossible to pin down in an extensional manner
4. *after the event*, it is often possible to pull out one or two key features of the context that “made all the difference.”

For reasons we have explained, however, the key to better intelligence analysis may not be to develop protocols or formalisms for making key context factors explicit and incorporating them into the reasoning process, as might naïvely be concluded from observation 4 in the above list — although such protocols or formalisms might be *part* of an overall solution. Making key contexts explicit can work — and has done so with great success — in a post mortem analysis, where the outcome is known and the causes are sought, but perhaps will not work for intelligence analysis, where the outcome is not known and the analyst is faced with gigabytes of potentially relevant information.

Perhaps then we should step back and ask whether our overall approach to dealing with context in intelligence analysis is the most appropriate one, given the goals of (i) developing technological aids to assist intelligence analysts in their work, and (ii) improving the process and its product.

The development of formal logics of context, which we alluded to earlier (such as McCarthy’s *ist* logic), is directed primarily at the first of these goals. (Clearly, however, progress toward goal (i) will yield progress toward goal (ii).) Our own earlier work described in [16], which we summarized in the previous section, is aimed simultaneously at both goals. (In trying to model the human intelligence analyst, we would measure progress in terms of new technologies, or of analytic processes and/or protocols, or of the education of new analysts.)

Both approaches begin with the assumption that we need to identify and formally represent the context in which a particular action takes place or in which a particular statement is valid. The work of Garfinkel and others tells us that neither approach can possibly “solve” the context problem, but that does not mean we cannot make considerable progress towards either or both of goals (i) and (ii).

On the other hand, it is not necessarily the case that progress toward either goal requires explicit representation of specific contexts. After all, people take account of context all the time when they engage in analysis and reasoning, yet rarely identify the specific contexts involved nor represent them explicitly in their reasoning process. Perhaps we should step back and look for an alternative approach.

Both previous approaches, formal context logics and our own framework, view the network shown in Figure 5 as one through which information flows. Both approaches effectively split the flowing information into two categories: (i) the target information, and (ii) the meta-information (pertaining to context) that allows the analyst to evaluate the reliability of the information and ultimately to decide what degree of trust to put in his conclusions.

Perhaps, a better approach (or at least a useful complementary one) is to view the flow through the network in Figure 5 not as one of information (including information about the analyst’s degree of trust in the source) *but of trust itself*. More generally, why not develop a framework for analyzing the way a network carries (in addition to information) such intentional entities as requests, commitments, commands, etc., the very entities that lead to the establishment of trust in the first place?

Such an approach amounts to a significant change of view and associated model, from a world-based, bottom-up, *commodity* approach (where the network supports the flow of stuff) to a human-agent-based, top-down, *intentional* one. In the former, data is the driver; in the latter the human drives. The reason we are able to make this shift and remain within the same overall framework is that the networks in both

Figure 5 and Figure 6 are two-sorted. There are two different kinds of entity being moved around the network. First, there are data, which are world-based. (See our analysis [15].) Then there are the human-based intentional entities such as information, concepts, promises, instructions, etc.

From a modeling viewpoint, one advantage of making such a change is the possibility of being able to make use of insights and methods from different domains, such as studies of commerce, particularly Internet commerce. For instance, consider the similarity between Figures 5 and

6.



Figure 6: The Internet: the view from 20,000ft.

In Figure 6, the individual sitting in the center is using the Internet to send and receive information, to convey instructions, to make requests, to make promises, to establish agreements, and to buy and sell goods. The electronic signals that flow through the network convey information, commitments, commands, and requests, leading ultimately to the establishment of trust. From 20,000ft, however, the activity of the Internet user in Figure 6 looks pretty much the same as the intelligence analyst in Figure 5. Thus, at a sufficient level of abstraction, we may expect that insights and results we obtain from a study of Internet use can carry over to the world of intelligence analysis. This is one of the primary benefits from abstraction, as we alluded to at the start of the previous section.

More generally, an abstractional approach of the kind we are advocating should lead to our being able to import ideas from several different domains, as indicated in

Figure 7.

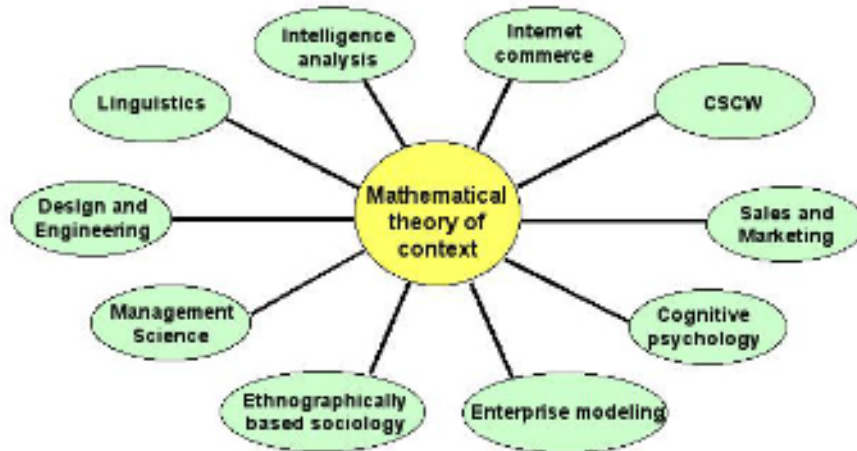


Figure 7: An abstract mathematical model of intelligence analysis can facilitate comparisons of the roles played by context in different domains, and perhaps lead to the transfer of solutions from one domain to another.

In particular, we see Figure 7 as the key to making progress on coming to grips with the six Ps of context. As we have already observed, humans and human organizations have evolved and developed various mechanisms to prevent problems of context — such as loss of common ground — from getting out of hand. A mathematical model of intelligence analysis at a suitable level of abstraction could help in identifying key mechanisms for coping with context effects in various domains of activity, and provide a channel whereby those mechanisms may be imported (with appropriate translation) from one domain into another — in particular (for our present purposes) into the domains of design and engineering and of intelligence analysis.

A particularly fruitful place to look for such “technology transfer” is work in psycho- and socio-linguistics, in both of which domains considerable work has been done on context.

For example, Clark and Carlson [12] talk about the *intrinsic context* of a discourse, the body of common knowledge shared by the participants on which the conversation is based. Clark [11] subsequently referred to this as the *common ground* of the discourse.

Del Hymes [27] proposed analyzing discourse by means of the SPEAKING model. This has eight components, whose first letters spell out the word “speaking”:

- **S**etting and scene. The setting is the time and place of the discourse, the scene is the cultural category, eg. a conversation, a lecture, or a speech.
- **P**articipants. The speaker, the audience, any overhearers.
- **E**nds. The purpose of the discourse, eg. a wedding speech, a plenary lecture.

- **Act sequence.** The format and order of the series of speech events that make up the speech act, eg. introduction followed by the body of the lecture followed by a Q&A session.
- **Key.** The cues that set the tone of the speech act. (The tone could be joking, somber, serious, etc. and the speaker could use dress or mannerisms to indicate the tone.)
- **Instrumentalities.** The forms and styles of speech used.
- **Norms.** The social rules governing the event and the participants' actions.
- **Genre.** Examples of genre are poem, myth, tale, proverb, riddle, prayer, oration, lecture, etc.

Harris [25] distinguished seven different *dimensions of context* for a discourse:

1. world knowledge
2. knowledge of language
3. authorial
4. generic (genre)
5. collective (general grasp of the relevant social institutions, customs, norms, etiquette, topical news items. etc.)
6. specific (specific to the actual discourse)
7. textual (the structure of the text, discourse, etc.)

The key to being able to take insights and discoveries from, say, sociolinguistics, and make use of them in technology-assisted intelligence analysis is to recast those discoveries in a suitable neutral, abstract framework, such as the one we are developing here. With this observation in mind, we should reconsider what kind of mathematical framework we shall adopt.

9 What kind of mathematics?

Although the research reported in this paper was carried out under subcontract to a major manufacturer of information systems, as part of a multi-year project to develop an information system to support intelligence analysis, the immediate focus of our own contribution, as reported in this paper, is not system design *per se*. Rather, we seek to develop a theoretical description of the entire intelligence analysis process — including both people (individuals and teams) and communication and information systems — grounded in an appropriate mathematical structure.

In particular, this means that our underlying mathematical framework must be able to represent human behaviors, system actions, and human-machine interaction, including the way contexts of various kinds influence those actions. This is about as wide a spectrum of requirements on a formal framework that one could imagine, particularly so for a mathematical framework. Certainly, mathematical formalisms such as first-order logic will not be adequate — at least if the intention is that the expressions in the formalism that we have to deal with can be read by humans.

One particular aspect of the uses to which such a framework will be put is to represent and analyze the kinds of human reasoning that cannot be effectively captured by first-order logic or logic-like formalisms (i.e., almost *all* kinds of human reasoning). Particular features of human reasoning that we wish to capture are:

1. It is not always linear.
2. It is often holistic.
3. The information on which the reasoning is based is often not known to be true. The reasoner must, as far as possible, ascertain and remember the source of the evidentiary information used and maintain an estimation of its likelihood of being reliable.¹¹
4. Reasoning often involves searching for information to support a particular step. This may involve looking deeper at an existing source or searching for an alternative source.
5. Reasoners often have to make decisions based on incomplete information.
6. Reasoners sometimes encounter and must decide between conflicting information.
7. Reasoning often involves the formulation of a hypothesis followed by a search for information that either confirms or denies that hypothesis.
8. Reasoning often requires backtracking and examining your assumptions.
9. Reasoners often make unconscious use of tacit knowledge, which they may be unable to articulate.

The factors we have just outlined imply that real-life evidence-based reasoning is rarely about establishing “the truth” about some state of affairs. Rather it is about marshaling evidence to arrive at a conclusion. If the reasoner wants to attach a reliable degree of confidence to the conclusion, she or he must keep track of the sources of all

¹¹Heuer [26, Chapter 4, p1] observes: “Judgment is an integral part of all intelligence analysis. While the optimal goal of intelligence collection is complete knowledge, this goal is seldom reached in practice. Almost by definition of the intelligence mission, intelligence issues involve considerable uncertainty. Thus, the analyst is commonly working with incomplete, ambiguous, and often contradictory data. The intelligence analyst’s function might be described as transcending the limits of incomplete information through the exercise of analytical judgment.”

the evidence used, the nature and reliability of those sources, and the reliability of the reasoning steps used in the process.

Concerning the requirement that the mathematical framework needs to be sufficiently flexible to express (i.e., represent) the key features of an analytic process *without distorting those features*, we note by way of self-caution that mathematics gains much of its strength by virtue of its simplistic approach; for example, complex entities become points in a space, complex relationships become surfaces or manifolds, etc. But this process of simplification can lead to a state of affairs where the mathematical results tell you a lot about the mathematical objects being studied and relatively little about the real-world entities they were intended to model. The analysis becomes largely theory internal.¹² This occurs when the mathematical properties of the model become more significant than the real world properties of the entities they were intended to model. It is a particularly prevalent danger with mathematical analyses because mathematicians, by inclination, training, and the culture of their discipline, constantly seek simplifying assumptions that lead to “better” and “more elegant” mathematics.

By good fortune, for the purposes of analyzing context effects in human decision making, a suitable mathematical framework already exists that (we believe) resists, if not completely mitigates, those dangers: situation theory.¹³ A particularly relevant and, we believe, highly beneficial, feature of situation theory is that it was developed, over many years, by mathematicians, linguists, and philosophers in order to analyze various deep and problematic phenomena in linguistics and human interaction through language (and not to provide a template for writing computer code).

Situation theory allows for an ontology tailored to a specific application domain, having entities of various kinds:

- individuals
- relations of different arities
- spatio-temporal locations
- types
- situations
- parameters

We shall make a few remarks about each of these entities momentarily. First though, we note that, with the exception of parameters, which are the theory’s mechanism for capturing informational links (and which are used as place-holders), all the objects in the ontology correspond to cognitive entities, i.e., entities individuated by a cognitive agent (the analyst or other humans in the case of intelligence analysis).

¹²This is exactly what happened with Catastrophe Theory, a branch of mathematics developed in the 1970s that never came close to meeting the original expectations of real-life applications.

¹³Had this not been the case, and were it not for the fact that we have been working with situation theory for many years, we would not have been involved with this project.

- **Individuals** are entities that the cognitive agent (and hence the theory) treats as single objects, ignoring their internal structure.
- **Relations** are viewed as real cognitive structures, and are not identified with sets of n -tuples (as is common practice in mathematics).
- **Locations** It would be possible to replace spatio-temporal locations by separate locations in space and time, and in fact this is often done in applications of situation theory.
- **Types** The ability to classify things in the environment by its type is fundamental for rational cognitive activity, which can be viewed as responding to a stimulus of one type by a response of another type.
- **Situations** are structured parts of the world. Examples of situations are a lecture, a telephone conversation, a video conference, the current military-political situation in Iraq. An obvious feature of situations is that in general they cannot be specified extensionally. Nevertheless, they have real cognitive existence, and are fundamental to rational behavior and information flow. It is largely by classifying by type the situations in which we continually find ourselves that we survive (if we do) and prosper (if we do). Much of military planning consists of envisaging and evaluating situations that do not yet exist.
- **Parameters** are in many ways the main workhorse of the mathematical apparatus that drives situation theory, much as the way that variables are central to real analysis. The early development of situation theory through the 1980s saw a great deal of discussion of whether parameters had any real significance outside the theory, or whether they were merely theoretical entities necessary to keep track of various informational — and other — links in the world that make rational activity and information flow possible. In discussions with Tomoyuki Yamada in the spring of 2004, we realized that Channel Theory, developed by Barwise and Seligman in the late 1990s and described in their book [9], provides a real-world meaning to parameters: parameters are path-types through channel-types.

Using a situation-theoretic ontology,¹⁴ it is possible to develop machinery for representing the way a token (such as a piece of text, some spoken language, or an action) in a context will encode (or generate) information that depends on both the token and the context. Figure 8 gives the general picture. (We shall see what a constraint is momentarily.)

The starting position is that, in an appropriate context, a sign or signal (which could be anything from a pile of stones or a knot in a handkerchief to a piece of text, a spoken sentence, a DNA molecule, or a 10 GB hard drive) can encode or carry or convey information. In situation theory, the sign or signal is represented as an individual or a

¹⁴Note that the ontology of situation theory is essentially an ontology *template*. Different application domains will require different actual ontologies.

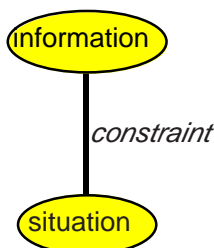


Figure 8: How signs or signals carry information.

situation and the context by a situation. The information is built up from basic (but not atomic) units called *infons*, which have the form

$$\ll R, a_1, \dots, a_n, l, t, i \gg$$

where R is an n -place relation, a_1, \dots, a_n are entities appropriate to fill the n argument roles in R , l is a location in space, t is a location in time, and i is a polarity, either equal to 1, in which case the infon is the item of information that the entities a_1, \dots, a_n are related by R at location l and time t , or else equal to 0, in which case the infon is the item of information that the entities a_1, \dots, a_n are not related by R at location l and time t .

In many (perhaps most) applications, the context situation will be too imprecise extensionally for a theorist to be able to carry out a sufficiently detailed analysis. To overcome this imprecision, we identify what exactly it is about the context that enables the sign/signal to carry the particular item of information it does. That crucial feature is provided (in situation theory) by a class of entities called constraints.¹⁵

Constraints are binary relations between situation types. We generally write a constraint in the form

$$S \Rightarrow T$$

This is read as “S involves T.”

Constraints may be natural laws, conventions, logical (i.e., analytic) rules, linguistic rules, empirical, law-like correspondences, etc. For example, humans and other agents are familiar with the constraint:

$$\textit{Smoke means fire.}$$

If S is the type of situations where there is smoke present, and S' is the type of situations where there is a fire, then an agent (e.g. a person) can pick up the information that there is a fire by observing that there is smoke (a type S situation) and being aware of, or *attuned to*, the constraint that links the two types of situation.

¹⁵Readers familiar with this term in other disciplines are best advised to ignore any other technical meaning they normally attach to this term and accept it simply as a word to which situation theory attaches its own specialized meaning.

This constraint is denoted by

$$S \Rightarrow S'$$

(This is read as “*S involves S’.*”)

Another example is provided by the constraint

FIRE *means fire.*

This constraint is written

$$S'' \Rightarrow S'$$

It links situations (of type S'') where someone yells the word FIRE to situations (of type S') where there is a fire.

Awareness of the constraint

FIRE *means fire*

involves knowing the meaning of the word FIRE and being familiar with the rules that govern the use of language.

The three types that occur in the above examples may be made precise as follows:

$$\begin{aligned} S &= [\dot{s} \mid \dot{s} \models \langle\langle \text{smokey}, \dot{t}, 1 \rangle\rangle] \\ S' &= [\dot{s} \mid \dot{s} \models \langle\langle \text{firey}, \dot{t}, 1 \rangle\rangle] \\ S'' &= [\dot{u} \mid \dot{u} \models \langle\langle \text{speaking}, \dot{a}, \dot{t}, 1 \rangle\rangle \wedge \langle\langle \text{utters}, \dot{a}, \text{fire}, \dot{t}, 1 \rangle\rangle] \end{aligned}$$

Notice that constraints link types, not situations. However, any particular instance where a constraint is utilized to make an inference or to govern/influence behavior will involve specific situations (of the relevant types). Constraints function by capturing various regularities across actual situations.

A constraint

$$C = [S \Rightarrow S']$$

allows an agent to make a logical inference, and hence facilitates information flow, as follows. First the agent must be able to discriminate the two types S and S' . (This use of the word ‘discriminate’ is not intended to convey more than the most basic of cognitive activities.) Second, the agent must be aware of, or behaviorally attuned to, the constraint. Then, when the agent finds itself in a situation s of type S , it knows that there must be a situation s' of type S' . We may depict this diagrammatically as follows:

$$\begin{array}{ccc} S & \Rightarrow & S' \\ s : S \uparrow & & \uparrow s' : S' \\ s & \rightarrow & s' \end{array}$$

For example, suppose $S \Rightarrow S'$ represents the constraint *smoke means fire*. Agent \mathcal{A} sees a situation s of type S . The constraint then enables \mathcal{A} to conclude correctly that there must in fact be a fire, that is, there must be a situation s' of type S' . (For this example, the constraint $S \Rightarrow S'$ is most likely reflexive, in that the situation s' will be the same as the encountered situation s .)

A particularly important feature of this analysis is that it separates clearly the two very different kinds of entity that are crucial to the creation and transmission of information: one the one hand the abstract types and the constraints that link them, and on the other hand the actual situations in the world that the agent either encounters or whose existence it infers.

We should point out that the ontology of situation theory has no bottom layer; every individual or situation can be subdivided into constituents, if desired. This implies that it is possible to represent and analyze a domain at any degree of granularity, to move smoothly up and down the granularity scale during an analysis, and to “zoom” the granularity to investigate specific issues in an analysis, while keeping the remainder of the representation fixed.¹⁶ We believe that this capability is extremely important for the kind of analysis we envisage being carried out using our framework.

We illustrate the way our framework may be applied in the study of context by describing two earlier projects we carried out with our colleague Duska Rosenberg, published in

[17], from which the following section is adapted.

10 Cultural contexts: an example

In his recent, important ethnographic study of the US intelligence analysis community, Johnston [28, p.6] stressed the major role played by the cultural context prevalent in the community:

“... risk aversion, organizational-historical context, and socialization are all part of the analytic process. One cannot separate the cognitive aspects of intelligence analysis from its cultural context.

Formal logic frameworks generally ignore cultural influences on reasoning, since culture is by and large not rule based. The framework we are developing, however, is well equipped to cope with culture, as we shall demonstrate by way of a seemingly simple example. (The apparent triviality of the example is highly misleading; it encompasses all the complexities of the way culture affects comprehension and reasoning.)

The example comes from the branch of sociology known as ethnomethodology, which sets out to analyze the minutiae of human action. In his seminal article [34], published in 1972, Harvey Sacks sought to illustrate the role played by social knowledge in our

¹⁶This feature played a major role in our analysis of engineer repair reports from a large computer manufacturer, described in [17].

everyday use of language. He took the following two sentences from the beginning of a child's story

The baby cried. The mommy picked it up.

and examined the way these two sentences are normally understood, paying particular attention to the role played by social knowledge in our interpretation of the story.

As Sacks observes, virtually every competent speaker of English understand this story the same way. In particular, we all hear it as referring to a very small human (though the word 'baby' has other meanings in everyday speech) and to that baby's mommy (though there is no genitive in the second sentence, and it is certainly consistent for the mommy to be some other child's mother). Moreover it is the baby that the mother picks up (though the 'it' in the second sentence could refer to some object other than the baby).

To continue, we are also likely to regard the second sentence as describing an action (the mommy picking up the baby) that follows, and is caused by, the action described by the first sentence (the baby crying), though there is no general rule to the effect the sentence order corresponds to temporal order or causality of events (though it often does so).

Moreover, we may form this interpretation without knowing what baby or what mommy is being talked of.

Why do we almost certainly, and without seeming to give the matter any thought, choose this particular interpretation? Sacks asks.

Having made all of his observations, Sacks explains [34, p.332]:

My reason for having gone through the observations I have so far made was to give you some sense, right off, of the fine power of a culture. It does not, so to speak, merely fill brains in roughly the same way, it fills them so that they are alike in fine detail. The sentences we are considering are after all rather minor, and yet all of you, or many of you, hear just what I said you heard, and many of us are quite unacquainted with each other. I am, then, dealing with something real and something finely powerful.

It is worth pausing at this point to emphasize our purpose in working through Sacks' example in some detail, as we shall do momentarily. After all, as Sacks himself notes, "the sentences we are considering are . . . rather minor." Yet, from the point of view of understanding the complexities of human interaction, the example embodies all of the key issues that arise. As Sacks observes, almost all of us understand the two sentences the same way. We do so despite the fact the practically none of that understanding is within the sentences themselves; it depends on our experience — what Sacks calls the 'fine power of a culture'.

One way to analyze the way the sentences are (normally) understood is to explicate the social relationships that are not overtly expressed. Sacks himself studied the

semantic strategies people use in communication. He showed how they may draw upon their knowledge of the social systems in order to arrive at shared interpretations of the actions they observe (or imagine, as in the case of the example of the child’s story). His main concern was to explain how shared social norms make such actions intelligible and interpretable (cf. [24, p.327]).

An alternative approach — which is the one Rosenberg and I adopted — is to identify the informational and cognitive *structures* that lead to the understanding, in particular the relational structures where relations that apply in a given situation represent the regularities the agent discriminates. The underlying structural form is indicated by the diagram on page 31. Our analysis has two main components. To identify which types S and S' are used and identify which constraints C connect those types. If we interpret Sacks own analysis in our framework (insofar as this is possible), we see that he formulates *rules* that explicate how the ‘fine power of a culture’ leads to the choice of types (our terminology) used to describe or understand the event or action. We use the type structure (i.e., the information-supporting structure) to explicate how that same ‘fine power of a culture’ guides the interpretation in a structural way.

Because the example, even though it may seem mundane, encompasses all of the main elements of human interaction, either form of analysis will result in insights and methods that have wide applicability.

In order to carry out our analysis, we need to introduce some situation-theoretic structures to represent the way that information flows from the speaker to the listener.

Reference to babies and mommies is captured in our framework by means of the types:

$$\begin{aligned} \text{‘baby’} &= T_{baby} = [\dot{p} \mid w \models \ll \text{baby}, \dot{p}, t_{now}, 1 \gg], \\ \text{‘mommy’} &= T_{mother} = [\dot{p} \mid w \models \ll \text{mother}, \dot{p}, t_{now}, 1 \gg], \end{aligned}$$

where \dot{p} is a parameter for a person. (In these type definitions, the situation w is “the world”, by which we mean any situation big enough to include everything under discussion. It is purely a convenience to think of this situation as the world, thereby providing a fixed context for the type definitions.)

We observe (as did Sacks in his original analysis) that both babies and mommies have different aspects. For instance, a baby can be thought of as a young person or as a member of a family, and a mommy can be viewed in relation to a child or to a father. These aspects, which affect the choice of words speakers make and the way listeners interpret them, are captured in our framework by the hierarchical structure on types (types of types, types of types of types, etc.).

Let:

$$\begin{aligned} T_{family} &= [\dot{e} \mid w \models \ll \text{family}, \dot{e}, t_{now}, 1 \gg], \\ T_{stage-of-life} &= [\dot{e} \mid w \models \ll \text{stage-of-life}, \dot{e}, t_{now}, 1 \gg], \end{aligned}$$

where \dot{e} is a parameter for a type.

The activity of crying is closely bound to babies in the stage-of-life type, so when the listener hears the sentence “The baby cried” he will understand it in such a way

that

$$(1) \quad T_{baby} : T_{stage-of-life}.$$

That is to say, this item of information will be available to the listener as he processes the incoming utterance, and will influence the way the input is interpreted.

Since the reader may be familiar to uses of “types” in other disciplines (such as computer science), where they are generally rigid in nature, we should stress that in situation theory, any type will typically be a member of an entire structure of types, and the applicability of a particular type may well depend upon two or more levels in the of-type hierarchy. For instance, the applicability of the type T_{baby} will be different when it is considered in the light of being in the type $T_{stage-of-life}$ as opposed to being in the type T_{family} . In the former case, individuals in the type T_{baby} will be typically and naturally associated with the activity of crying (type T_{crying}); in the latter case they will be typically and naturally associated with having a mother (2-type $T_{mother-of}$). (In situation-theoretic terms, these associations will be captured by constraints that link types. Those constraints are in general not universals, rather they may depend on, say, individual or cultural factors.) This particular distinction will play a significant role in the analysis that follows.

One immediate question concerns the use of the definite noun phrases ‘the baby’ and ‘the mommy’. Use of the definite article generally entails uniqueness of the referent. In the case of the phrase ‘the baby’, where, as in the Sacks example, no baby has previously been introduced, one would normally expect this to be part of a more complex descriptive phrase, such as ‘the baby of the duchess’s maid’, or ‘the baby on last night’s midnight movie’. So just what is it that enables the speaker to open an explanation with the sentence ‘The baby cried’? It could be argued that an implicit suggestion for an answer lies in his later discussion of proper openings for ‘stories’, but this is a part of his article we do not consider here.

For a situation-theoretic analysis, there is no problem here. The situation theorist assumes that all communicative acts involve a *described situation*, that part of the world the act is *about*. Exactly how this described situation is determined varies very much from case to case. For example, the speaker may have witnessed, read about, or even imagined the event she describes. In the Sacks example, the speaker *imagines* a situation in which a baby cried and its mother picked it up. Let s denote that situation.¹⁷

The situation s will be such that it involves one and only one baby, otherwise the use of the phrase ‘the baby’ would not be appropriate. In starting a communicative act with the sentence ‘The baby cried’, the speaker is informing the listener that she is commencing a description of a situation, s , in which there is exactly one baby, call it b . (Whether or not b is a real individual in the world, or some fictional entity, depends

¹⁷It does not affect the mechanics of our analysis whether you think of situations as objects in the speaker and listener’s realm — possibly as things they are aware of — or purely as theorist’s objects in an abstract ontology adopted to study interaction. All we need to know is that these situations are definite objects available to the theorist as part of a framework for looking at the world. In the case where situations are regarded purely as theorist’s abstractions, s will *correspond* to some feature of the interaction—you can think of s as providing us with a *name* for that feature.

on s . This does not affect the way our analysis proceeds, nor indeed the way people understand the utterance.)

The principal item of information about the described situation that is conveyed by the utterance of the first sentence ‘The baby cried’ is

$$s \models \ll \text{cries}, b, t_0, 1 \gg$$

where t_0 is the time, prior to the time of utterance, at which the crying took place. In words, in the situation s , the baby b was crying at the time t_0 .

Notice that, in the absence of any additional information, the only means available to the listener to identify b is as the referent for the utterance of the phrase ‘the baby’. The utterance of this phrase tells the listener two pertinent things about s and b :

$$(2) \quad b : T_{baby} \quad (\text{i.e. } b \text{ is of type } T_{baby})$$

where T_{baby} is the type of all babies, and

$$(3) \quad b \text{ is the unique individual of this type in } s.$$

Now let’s consider what additional information is conveyed by the utterance of second sentence, ‘The mommy picked it up.’ Mention of both babies and mommies invokes the family type, T_{family} . This has the following structural components that are relevant to our analysis:

| | |
|-----------------|--|
| $M(x)$ | the property of x being a mother |
| $B(x)$ | the property of x being a baby |
| $M(x, y)$ | the relation of x being the mother of y |
| T_{mother} | the type of being a mother |
| T_{baby} | the type of being a baby |
| $T_{mother-of}$ | the 2-type that relates mothers to their offspring |

In the type T_{family} , the type $T_{mother-of}$ acts as a fundamental one, with the types T_{mother} and T_{baby} being linked to, and potentially derivative on, that type. More precisely, the following structural constraints¹⁸ are salient in the device T_{family} :

$$\begin{aligned} T_{mother} &\Rightarrow \exists y T_{mother-of} \\ T_{baby} &\Rightarrow \exists x T_{mother-of} \end{aligned}$$

where

$$T_{mother} = [\dot{x}, \dot{y} \mid w \models \ll \text{mother-of}, \dot{x}, \dot{y}, t_{now}, 1 \gg].$$

What do these mean? Well, $T_{mother-of}$ is a 2-type, the type of all pairs of individuals x, y such that x is the mother of y (at the present time, in the world). The first of

¹⁸The notion of constraint used here extends that described in [?].

the above two constraints says that the type T_{mother} *involves* (or is *linked to*) the type $\exists y T_{mother-of}$. This has the following consequence: in the case where $T_{mother} : T_{family}$ (i.e. T_{mother} is of type T_{family}) and $T_{baby} : T_{family}$, the following implications are salient:

$$(4) \quad p : T_{mother} \rightarrow \exists q (p, q : T_{mother-of})$$

$$(5) \quad q : T_{baby} \rightarrow \exists p (p, q : T_{mother-of}).$$

These two implications are not constraints. In fact they do not have any formal significance in situation theory. They are purely guides to the reader as to where this is all leading. (4) says that if p is of type T_{mother} (i.e. if p is a mother), then there is an individual q such that the pair p, q is of type $T_{mother-of}$ (i.e. such that p is the mother of q). The salience of this implication for an agent \mathcal{A} has the consequence that, if \mathcal{A} recognizes that p is a mother then \mathcal{A} will, if possible, look for an individual q of which p is the mother. Analogously for (5).

To continue with our analysis, as in the case of ‘the baby’, in order for the speaker to make appropriate and informative use of the phrase ‘the mommy’, the described situation s must contain exactly one individual m who is a mother. In fact we can make a stronger claim: the individual m is the mother of the baby b referred to in the first sentence. For if m were the mother not of b but of some other baby, then the appropriate form of reference would be ‘a mother’, even in the case were m was the unique mother in s . We can describe the mechanism that produces this interpretation as follows.

Having heard the phrase ‘the baby’ in the first sentence and ‘the mommy’ in the second, the following two items of information are salient to the listener:

$$(6) \quad m : T_{mother}$$

$$(7) \quad m \text{ is the unique individual of this type in } s.$$

In addition, we shall show that the following, third item of information is also salient:

$$(8) \quad m \text{ is the mother of } b.$$

Following the utterance of the first sentence, the listener’s cognitive state is such that the type T_{baby} is of type $T_{stage-of-life}$. This type has categories that include T_{baby} , T_{child} , $T_{adolescent}$, T_{adult} , all of which have equal ontological status within the stage-of-life type, with none being derivative on any other. But as soon as the phrase ‘the mommy’ is heard, the combination of ‘baby’ and ‘mommy’ switches the emphasis from the type $T_{stage-of-life}$ to the type T_{family} , making salient the following propositions:

$$(9) \quad T_{baby} : T_{family}.$$

$$(10) \quad T_{mommy} : T_{family}.$$

In the T_{family} device, the various family relationships that bind a family together (and which therefore serve to give this type its status as a type) are more fundamental

than the categories they give rise to. In particular, the types T_{baby} and T_{mother} are derivative on the type $T_{mother-of}$ that relates mothers to their babies.

Now, proposition (9) is the precondition for the salience of implication (5), namely

$$q : T_{baby} \rightarrow \exists p(p, q : T_{mother-of}).$$

Substituting the particular individual b for the variable q , we get

$$b : T_{baby} \rightarrow \exists p(p, b : T_{mother-of}).$$

But by (2), we know that

$$b : T_{baby}.$$

Thus we have the salient information

$$(11) \quad \text{there is an } m \text{ such that } m, b : T_{mother-of}.$$

The use of the definite article in the phrase ‘the mommy’ then makes it natural to take this phrase to refer to the unique m that satisfies (11). Thus the listener naturally takes the phrase ‘the mommy’ to refer to the baby’s mother. This interpretation is reinforced by the completion of the second sentence ‘...picked it up’, since there is a social norm to the effect that a mother picks up and comforts her crying baby. This explains how the fact (8) becomes salient to the listener.

It should be noticed that the switch from the salience of one set of constraints to another was caused by the second level of types in the hierarchy. The constraints we were primarily interested in concerned the types T_{mother} and T_{baby} . These types are part of a complex network of inter-relationships (constraints). Just which constraints in this network are salient to the agent is governed by the way the agent encounters the types, that is to say, by the type(s) of those types—for instance, whether T_{baby} is regarded (or encountered) as of type $T_{stage-of-life}$ or of type T_{family} . By moving to a second level of typing (i.e. to types of types), we are able to track the way agents may use one set of constraints rather than another, and switch from one set to another. The first level of types allows us to capture the informational connections between two objects; the second level allows us to capture the agent’s preference of a particular informational connection, and thereby provides a formal mechanism for describing normality. This helps to bridge the gap between two distinct traditions, each with a different take on context: the descriptive, which focuses on the particulars, and the formal, focused on the universal.

As we indicated earlier, the fundamental nature of the issues embodied in the Sacks example means that the methods we employed in our analysis have much wider applicability. For instance, in the late 1980s and early 1990s, Rosenberg and I analyzed what had gone wrong when a large manufacturer and supplier of mainframe computer systems had tried to automate part of its own information system, namely the data collected in the standard form (the Problem Report Form, or PRF) filled in when an engineer was called out on a repair job.

The PRF was a simple slot-and-filler document on which could be entered various reference numbers to identify the customer and the installed system, the fault as reported by the customer, the date of the report, the date of the engineer's visit, the repair action he took, and any components he replaced.

The PRF was a valuable document, providing the company with an excellent way to track the performance of both their computer systems and their field engineers, as well as the demand for spare parts. In particular, by analyzing the data supplied by the forms, the company could identify and hopefully rectify the weakest components in the design of their systems.

Because of the highly constrained nature of the PRFs, the highly focused nature of the domain — computer fault reporting and repair — and the fact that the important technical information on the forms was all entered by trained computer engineers, the company had expected that the PRFs would form the basis of a highly efficient source of information for all parts of the company. But that was not how it turned out. Both in the early days, when the PRFs were paper documents, and later when they were replaced by an electronic version, experts faced with reading the forms frequently encountered great difficulty understanding exactly what had gone wrong with the customer's system and what the engineer had done to put it right.

The problem was magnified when the company tried to develop an expert system to handle the information provided by the PRFs. An initial investigation led to the suspicion that the problem was caused at least in part by the fact that understanding entries on the form often required social and/or cultural knowledge familiar to only one segment of the company's personnel (and not familiar at all to the expert system).

Applying extensions of the techniques used to analyze the Sacks example, we were able to carry out a detailed analysis of the way social and cultural knowledge affected the information conveyed by the PRFs. This led to a restructuring of the procedures surrounding the completion and use of the documents, resulting in better information flow and improved efficiency in the company.

Furthermore, the additional problem our analysis addressed was to relate the structure of the document to its broader uses in the organization as a whole. We viewed the PRF as a resource that facilitates (or obstructs, as the case may be) the interaction between different sections of the organization. In this context, the social significance of the document needs to be understood so that the information flow between different sections may be organized and managed.

An investigation into the uses of the document, as opposed to its structure, brought to light the need to develop a dual perspective — document intension and schema of investigation, illustrated in Figure 9. The document intension captures the communicative intent of the various sections of the document, through the use of the constraints that formalize the informational links within the document. The schema of investigation traces the information pathways a reader of the document creates in the process of interpretation. This is schematically presented in Figure 10.

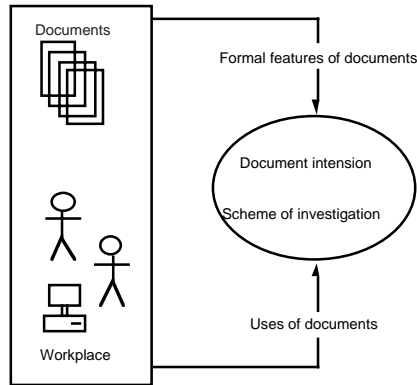


Figure 9: Document intension

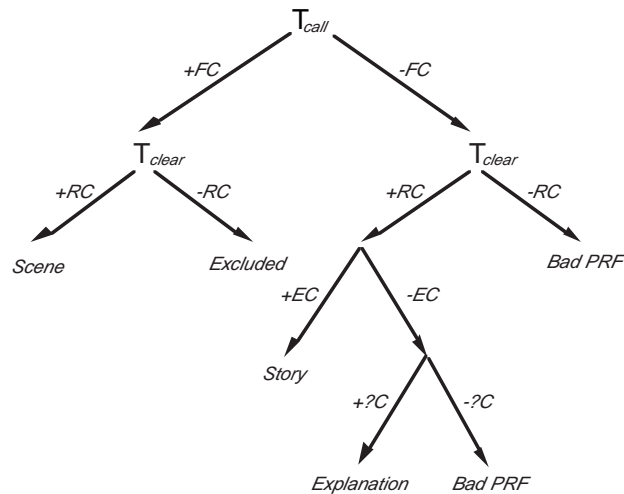


Figure 10: Interpretive grammar

The schema captures formally how the successive application of constraints leads to “perfect” information in the “scene”, when everything fits — on the far left of the tree — and also to the “bad PRF” on the far right of the tree. These examples illustrate the strategies that computerized resources capture easily.

However, most of the everyday cases analyzed were not so clear cut. Going from left to right in the tree in Figure 10, if the fault description is clear and the remedial action is not, this would be interpreted as the engineer not knowing his job. Needless to say, no PRF among the hundreds analyzed gave this information explicitly. The most frequent and the most challenging examples were those in the middle of the tree, where the fit had to be established between the fault description, the appropriate remedial action and the resources used in implementing the remedy. This is where most of human interpretive effort was focused. Sadly, this is also where computerized tools are

still grossly inadequate as they are not responsive to the human uses of the information stored in them.

An empirical study the uses of the PRF in the organization showed that the information contained in the document was needed to support the decisions of customer services, fault diagnosis, spare parts and change management, and production. On a more strategic level, the information flow mediated by the PRF was used as the basis for evaluating the effectiveness of various organizational processes that eventually led to an award-winning reorganization of after-sales maintenance.

We should stress that, even though both examples presented in this section, the baby and the fault record, concern particular interactions in particular contexts, the analytic methods used to analyze them, described above, are able to capture the underlying regularities, or uniformities, and hence can be generally applied.

11 Extending the Barwise–Perry agent model

We now return to the goal of developing a mathematical analysis of the way that commands, promises, requests, and trust flow around a network.

In their initial development of situation semantics, as described in their book *Situations and Attitudes* [4], Barwise and Perry modeled the act of one person uttering a declarative sentence (or a part thereof) to another by (effectively) treating the two people as points. They focused their analysis on a small number of abstract *connections* between those points and the surrounding world, a small number of parts of the world (*situations*), a small number of abstractions called *types*, some abstract connections between types (*constraints*), and some framework-internal technical machinery. They ignored practically every aspect of the two people, their ages, heights, weights, skin colors, histories, personalities, achievements, professions, and so on. (Occasionally their model did represent names and their genders.) In particular — and this is the truly significant decision they made in setting up their model — they did not seek to represent the cognitive states of the two individuals. Even when they were analyzing attitude reports (such as “Jon sees that . . .”), they did not try to “get inside the minds” of the individuals being modeled.

That choice — for it was a very conscious and definite choice — led some critics to accuse them of engaging in, or at the very least flirting with, behaviorism, but to make such an accusation is to misunderstand in a fundamental way how mathematical modeling works. The most important aspect of any mathematical model is the initial choice of what to represent in the model — and what to leave out. Mathematical models typically ignore almost every feature of the thing being modeled. (They also idealize the small number of features that are represented in the model, but this is secondary, because the idealization can only occur after the initial choice of features has been made.)

For example, the Solar System may be (and often is) usefully modeled as a system of point masses with one designated the “center” (representing the Sun) and the others

orbiting around it in elliptical orbits. Only the mass of the Sun and the masses, relative positions, and relative velocities of the individual planets are represented. The sizes, exact shapes, chemical compositions, atmospheres, temperatures, colors, ages, histories, etc. of the Sun and the planets are all ignored. When Galileo and Kepler and later Newton analyzed the motion of the planets in terms of the point masses model we referred to above, no one, surely, accused them of believing, or claiming, that the Sun and the planets *are* just point masses. Rather, it was understood that they simply took them to be point masses for the purposes of the model. Mathematicians at NASA make exactly the same assumptions today, with great effect, when they send rockets to explore the Solar System. This approach worked, and still works, because the orbital behavior of the system is what matters when it comes to intended applications. Similarly, Barwise and Perry elected to focus on a small number of “behavioral aspects” of linguistic exchanges in order to develop their mathematical model.

Remarkably, although we don’t recall anyone ever noting just how remarkable this was, their approach served to provide a useful analysis of various kinds of linguistic exchanges — cognitive activities if ever there were — including *sees that* attitude reports. It was only when they came to belief reports that they acknowledged a need to represent (in the model) cognitive states, or at least to consider types that classify such states. Similarly, when we analyzed the Searle speech act taxonomy¹⁹ in our book *Logic and Information* [13, Section 8.5], we did so using types that classified cognitive states.

For both Barwise–Perry and for ourselves, a natural next step would have been to analyze further those cognitive types — those types of certain cognitive states — but we did not do so. In our case, not because we thought it was not an important issue; rather, other questions took our attention. We suspect the same was also true for Barwise and Perry.

At this stage, we need to be more specific about the Barwise–Perry model, since to the best of our knowledge, in all previous studies (including both theirs and our own) everyone left this implicit.

For the Barwise–Perry analysis, a conversational agent is modeled as:

1. an individual a (atomic, unanalyzed);
2. a collection \mathcal{H}_a of types (a ’s current classification scheme);
3. a collection \mathcal{I}_a of infons (the information currently possessed by a);
4. a mechanism (binary relation) OF-TYPE_a whereby the agent ascribes types to things (objects, other agents, actions, situations, etc.);
5. a mechanism (binary relation) REFERS-TO_a whereby the agent links certain linguistic utterances to things in the world (speaker’s connections, referents of proper names and pronouns, reference to the described situation, etc.).

¹⁹Assertives, directives, commissives, declarators, and expressives. See Searle [35].

We shall call a structure of the kind just specified a *Barwise–Perry model* (of a conversational agent).

Our proposed new model (which we shall modestly name a *Devlin model*) adds one further feature:

6. a collection \mathcal{C}_a of personal constraints.

The constraints in \mathcal{C}_a govern the actions of the individual a based on his or her current knowledge, beliefs, desires, and intentions. As such, they are the external (or behavioral, and hence observable) manifestations of the agent’s knowledge, beliefs, desires, and intentions.²⁰

We believe that this enriched model will facilitate a powerful extension of situation semantics — which focuses on the transmission of information — to an action-oriented analysis of language use. But that awaits further work. In the meantime, notice that, from a mathematical perspective, moving from the Barwise–Perry model to our new one amounts to a shift from working in *information space* (whose elements are infons) to *constraint space* (whose elements are constraints). This is analogous to the familiar step in mathematics from a given space to its function space. We intend to investigate this aspect of our new model in a subsequent paper.

For now, however, our next step is to see how our new model provides a simple and straightforward classification of the basic actions that make up interpersonal and commercial Internet activity.

12 Modeling basic Internet actions

A basic Internet transaction occurs when one person or computer sends an electronic message to another. In order to draw useful comparisons with situation semantics, which focuses on one person speaking to another, it will be helpful to cast our initial discussion in terms of “speakers”, “listeners”, and “utterances”. Thus, the person or computer that sends the message will be called the “speaker”, the receiving agent will be the “listener”, and the message will be the “utterance”.

In contrast to situation semantics, which concentrates on the information conveyed by a linguistic act, we seek to examine the effect a speaker wishes to achieve by a particular utterance and the effect it actually has. Within our new model, the intended and actual effects are captured by changes to the constraint sets \mathcal{C}_a and \mathcal{C}_b of speaker and listener. In order to provide a brief preview of where we are heading, let’s see how the new model differentiates various kinds of speech acts.

Suppose a utters a sentence Φ to b . Then:²¹

²⁰In this note we shall focus almost entirely on intentions.

²¹The terminology we use has considerable overlap with that of Searle [35], but our usage is not exactly the same as his. He was considering illocutionary acts, whereas our focus is more on perlocution. However, in due course we shall in any case have occasion to look more closely at what constitutes a basic perlocutionary act, so for now we’ll simply coopt the terminology for our own purposes. Note

- The utterance of Φ is an *assertive* if a primarily intends the utterance to modify \mathcal{I}_b .
- The utterance of Φ is a *directive* if a primarily intends the utterance to modify \mathcal{C}_b and satisfaction of the act depends on b 's subsequent action.
- The utterance of Φ is a *commissive* if a primarily intends the utterance to modify \mathcal{C}_a or to make public to b a constituent of \mathcal{C}_a , and satisfaction of the act depends on a 's subsequent action.
- The utterance of Φ is a *declarator* if a primarily intends the utterance to modify the community facts and/or the collection of shared constraints that govern the actions of a particular community.
- The utterance of Φ is an *expressive* if a does not intend the utterance to modify any personal or community information or constraints.

Note that this classification does not require that we say which collections are modified, or even what kinds of constraints are modified, if any. It is enough to identify the collection the speaker intends to modify: its type — information or constraints — and in the latter case whether it is the speaker's, the listener's, or the community's constraints.

A comment is in order regarding our use of the word “primary” in the above. When a speaker utters a directive or a commissive, the purpose is to bring about a change in the listener. In the case of a directive, that change is the formation by the listener of an intention to act as instructed. For a commissive, however, things are a bit more complex. The speaker who utters a commissive is declaring an intention to act a certain way. The speaker's subsequent actions will determine whether or not the commissive is satisfied. For that reason, we say the *primary* intention is to form or make public to the listener the intention concerned. But if everything is up to the speaker, why make the utterance in the first place? Why not keep the intention private and simply act upon it at the appropriate time? The answer, of course, is that we make promises in order to ensure cooperation. By making the promise, the speaker creates expectations in the listener and entitles the listener to act in certain ways. Since this is the reason for uttering a commissive, it could be argued that the primary intention is the effect on \mathcal{C}_b , not that on \mathcal{C}_a . Our use of the word “primary” in this case is at odds with this observation, and is chosen to provide a consistent terminology across the different (perlocutionary) speech acts. In other words, our choice of terminology on this occasion reflects the framework within which we carry out our analysis, rather than the domain of application. According to our terminology, the primary intention is the one related to the constraint that must be satisfied in order to provide compliance with the speech act.

In fact, even in the case of a directive, there is an effect on the speaker's constraint set, since issuing a command puts expectations and obligations on the speaker. Thus,

that the criteria we give are not meant to provide iron-tight *if and only if* definitions; rather they give distinguishing characteristics of the different kinds of perlocutionary acts.

there is more symmetry between commissives and directives than might at first appear, albeit of different strengths. We shall examine these issues in more depth later.

An interesting special case of commissives arises in the case of suicide bombers, where the act is almost always preceded by a ritualistic ceremony at which the bomber makes a public commitment to carry out the act. In this case the purpose is entirely to impose sociological constraints on the perpetrator that make it psychologically impossible not to carry through on the commissive.

The new model allows us to make the distinction between an agent having information and having knowledge. To say that *a* has a particular item of information σ is simply to say that $\sigma \in \mathcal{I}_a$. If *a* knows σ , then there will be a corresponding constraint or collection of constraints in \mathcal{C}_a . Barwise and Perry certainly discussed the distinction between information and knowledge, as we did in [13] and elsewhere, and in both studies the essence of the distinction is that knowledge is information that the agent has internalized in a form that can influence action. But in neither treatment was the distinction captured directly at the level of the model, as it is here.

We shall investigate the above classification in more depth later in a later paper. For now, some further remarks are in order concerning mathematical modeling.

13 The importance of models

Choosing which features of a domain to incorporate in a mathematical model is much more than a mere technical matter that affects only what mathematics gets done. If the model turns out to have recognizable analytic or predictive power, and in particular if it also provides a simplified picture of a complex domain that others can use, it can have a fundamental influence on the way nonmathematical domain experts approach the field. Perhaps the most dramatic illustration of this is the “miniature solar system” model of the atom that Niels Bohr put forward in the early twentieth century. Scarcely twenty years passed before physicists recognized that the Bohr model was not at all a realistic image of the atom. Nevertheless, in the absence of an equally simple and compelling alternative model, the Bohr model lives on to this day, and influences the way we think of atoms. It even forms the logo of the International Atomic Energy Agency, a body whose leaders and learned members know full well that the atom isn’t like that.²²

A similar critique could be made of the model of the human brain as a digital computer, but since many leading and learned members of the cognitive science community continue to believe in that picture — or so it seems from their writings — it may be too great a digression for me to pursue such a critique here.

The Barwise–Perry model of conversational agents — as we observed before, implicit in all their work on situation semantics, even if they never spelled it out as a model

²²Of course, part of the reason the Bohr model lives on is that, despite its inaccuracy, it remains extremely useful. A lot of very good, and often practical chemistry, biology, and even physics can be done using that model as the basis.

— was extremely useful, in particular for analyzing definite descriptions, indexicals, anaphora, and certain attitude reports.

The new agent model we are proposing here may look like a small perturbation of their model, but it is far more than that. By including the agent’s behavior-influencing constraints in the model, we take the theory from one of information content to one of action — a major shift. The fundamental question that the Barwise–Perry model addressed was “What information is transmitted from one person to another by an utterance?” With our new model, the fundamental question is “What effect does an utterance by one person have on the beliefs, desires, intentions, or actions of the addressee?” Because the Barwise–Perry model was developed to handle information flow, those were the kinds of questions researchers asked of it. Given a model built on action and the way linguistic utterances function as actions and affect actions, the kinds of questions that will be asked when our new model is assumed will be about action. As the saying goes, when the only tool you have is a hammer, everything starts to look like a nail. Our reason for proposing the new agent model is that, according to many who work in semantics, the action nail is where the future action in the field is likely to be found.

Notice that our new model does not seek to represent cognitive states. Rather, we incorporate the agent’s *constraints* that result from her or his current cognitive state. Thus, our model remains external to the agent, just as did situation semantics.

Since the only thing we have added to the Barwise–Perry model is the constraint set, when we use our new model (to study, say, Internet commerce), we do not have to start all over again. Instead, we may make use of all of the machinery developed by the situation semantic and situation theory community over the years, in particular the extensive apparatus for handling constraints themselves.

One particular advantage of working with existing machinery is that we can often avoid having to resort to heavy use of mathematical formalisms, knowing from previous work that the formalism can be provided if desired. Thus, we can focus much more on the main issues, and avoid getting bogged down in questions about the underlying mathematical apparatus, a problem that plagued Barwise and Perry for most of the early stages of their work. (Well, at least that’s our hope!)

14 Perlocutionary acts and intentions

Situation semantics focuses on the information conveyed by a linguistic utterance. But language is used for purposes other than the transmission of information, important though that usage is. Particularly important uses of language that can be handled only partially using situation semantics are to give commands, make commitments, and issue requests. Such uses of language are called *illocutionary* acts. Their successful completion (successful in the sense of the speaker’s intention being achieved) are called *perlocutionary* acts. Our new model of a conversational agent is designed to permit the analysis of perlocutionary uses of language by way of the utterance of illocutionary

acts. First, the general idea.

When person A issues a command to person B , A intends that B forms an intention to act as A commands. If B does in fact form such an intention, and moreover does so because of A 's command, we say that B *accepts* the command. If, in addition, B actually carries out A 's instruction (and does so because so instructed), we say that B *meets* A 's command.

Similarly, suppose person A issues a commissive to person B . To say that A utters the commissive *faithfully* is to say that A forms an intention to act according to the promise. If A actually carries out the promised action, and does so by virtue of having made the promise, we say that A *keeps* the promise.

Finally, at the level of our current analysis, we shall regard a request as a polite form of command, and hence our previous treatment of commands applies also to requests. Thus, if A makes a request of B , we say B *accepts* the request if he or she forms an intention to do as A requests, and moreover does so as a result of A 's request. If, in addition, B realizes that intention and thereby carries out the request, we say he or she *meets* A 's request.

We recognize that in treating requests as polite commands, we are ignoring a number of issues. In particular, whereas the recipient of a command has just two options, to obey or not, a broader range of compliant responses is available to the recipient of a request. We intend to look further into this matter at a later date, when we have the mathematical machinery more fully worked out.

In each of the two cases we have considered, commands/requests and commissives, we gave two “success criteria,” the first in terms of forming an intention and the second performing an action that fulfills that intention. We do this in order to capture the distinction — which we feel is important — between the speech act *qua an act in itself* and the action the speech act is about. For instance, a command succeeds as a speech act provided the commandee accepts the instruction and forms an intention to do as instructed. If events subsequently transpire that prevent the commandee from carrying out that intention, that does not nullify the successful completion of the speech act.

We also imposed conditions that exclude the incidental (or accidental) “satisfaction” of a speech act, demanding that the intention is formed as a result of the speech act, and that any subsequent act that accords with the intention is a result of the intention being fulfilled. The essence of a speech act is that it brings about a specific intended result — in the first place the formation of an intention to perform some action and ideally the subsequent performance of the specified action. If that action is subsequently performed *for some other reason*, then both parties may well feel satisfied with the outcome, but the result will not have been achieved because of the speech act, and hence should not be regarded as a genuine fulfillment of the speech act.

In order to formalize these ideas, we shall develop machinery to represent intentions as situation-theoretic constraints. We shall do this first in a way that captures most, but not all, of our intuitions. Then we shall modify our approach to eliminate the gap.

Consider first the case where A issues the command to person B :

IF ϕ , DO ψ .

(We will regard commands of the simpler form DO ψ as special cases of the more general framework.) This will determine a constraint $T_1 \Rightarrow T_2$. We will call T_1 , which is related to ϕ , the *prior type*, and T_2 , related to ψ , the *action type*. The interpretation of A 's utterance of ϕ is $u : T_1$, where the situation u will be given by the speaker's connections, and the interpretation of A 's utterance of ψ is $v : T_2$, for an appropriate situation v . In short, if the specified situation u is of type T_1 , then B brings it about that a certain situation v is of type T_2 .

Here is the formal development.

By an *action type* of an agent a we mean a situation type of the form

$$T = [\dot{u} \mid \dot{u} \models \langle\langle a, W, \dots, \dot{i}_{here}, \dot{t}_{now}, 1 \rangle\rangle]$$

where W is an action.

By an *intention* of a we mean a constraint of the form

$$C = T_1 \Rightarrow T_2$$

where T_1 is a situation type, called the *prior* of the intention, and T_2 is an action type of a , called the *completion*.

A pair (u, v) of situations is said to *satisfy* C if $u : T_1$ and $v : T_2$.

Given a situation u , we say that C is *fulfilled with respect to* u if there is a situation v such that

- (i) (u, v) satisfies C
- (ii) $v : T_2$ by virtue of C

We say that the intention C is *fulfilled* if there is a pair (u, v) of situations such that v is fulfilled with respect to u .

A command K given by A to B is said to be *accepted* if the associated intention C is incorporated into \mathcal{C}_B . The command is *met* if there is a situation v that fulfills C with respect to u , where u is determined by the speaker's connections function (given by the demonstrative conventions).

For example, suppose that Alex (A) says to Betty (B):

IF GEORGE IS AT THE PARTY, GIVE HIM THIS RING.

By the speaker's connections, there is an individual G whom Alex refers to by the word GEORGE, a situation p referred to by the phrase THE PARTY, and an object r referred to by the phrase THIS RING. Then the constraint determined by Alex's utterance is $C = T_1 \Rightarrow T_2$, where:

$$\begin{aligned} T_1 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{present}, G, \dot{i}_{here}, \dot{t}_{now}, 1 \rangle\rangle] \\ T_2 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{gives}, B, G, r, \dot{i}_{here}, \dot{t}_{now}, 1 \rangle\rangle] \end{aligned}$$

Alex’s command is accepted if the constraint C is added to \mathcal{C}_B (by B), and is met if there is a situation v that fulfills C with respect to p . (If you assume that Alex’s command has the implicature that Betty should give George the ring at the party, then the command will be met if p itself fulfills C with respect to p .)

For another example, suppose Alex says to Betty:

YOU MUST LEAVE IMMEDIATELY.

By the speaker’s connections, there is a physical location (a situation) p , implicitly referred to by A , which B is instructed to leave. This time the constraint determined by Alex’s utterance is $C = T_1 \Rightarrow T_2$, where:

$$\begin{aligned} T_1 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{present}, B, \dot{t}_{now}, 1 \rangle\rangle] \\ T_2 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{present}, B, \dot{t}_{now} + \epsilon, 0 \rangle\rangle] \end{aligned}$$

Alex’s command is accepted if the constraint C is added to \mathcal{C}_B (by B), and is met if p (itself) fulfills C with respect to p .

Commissives are handled in an entirely parallel fashion. A commissive K given by A to B is said to be *faithful* if the associated intention C is either in or else is incorporated into \mathcal{C}_A . It is *kept* if there is a situation v that fulfills C with respect to u , where u is determined by the speaker’s connections function.

For example, suppose presidential candidate Shrub (S) makes the promise to the electorate of country Gullible (G):

IF ELECTED, I WILL LOWER TAXES.

The constraint determined by Shrub’s utterance is $C = T_1 \Rightarrow T_2$, where:

$$\begin{aligned} T_1 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{elected}, S, l_G, \dot{t}_{election}, 1 \rangle\rangle] \\ T_2 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{lower-taxes}, S, l_G, \dot{t}_{now}, 1 \rangle\rangle] \end{aligned}$$

If Shrub’s promise is faithful, then either the constraint C is in \mathcal{C}_S (and his utterance makes it public) or else he adds the constraint C to \mathcal{C}_S . If there is a situation u such that $u : T_1$ (i.e., if Shrub is elected), then the promise will be kept if there is a situation v that fulfills C with respect to u .

Since we are treating requests as just polite forms of commands, that completes our examples, and it’s time to plug the gap we said was in our account so far. The issue of concern is the distinction between an action that is carried out to meet the requirements created by a command, commissive, or request, and an action that meets those requirements only incidentally. As it stands so far, our framework does not fully capture this distinction, since we have not provided any direct link between the performance of a speech act and the subsequent performance of the appropriate action. Rather, what our framework requires is that the performance of the action is carried out in order to follow the specified constraint. Our idea is that the constraint arises

by virtue of the speech act, but the existing machinery of situation theory does not provide a mechanism to track the origins of constraints. In earlier work on situation semantics, this was not required. But to develop a situation-theoretic treatment of speech acts, we need to capture such connections. This is why we postponed this step until now.

If a constraint $C = T_1 \Rightarrow T_2$ is created or made public by virtue of a speech act u (a situation), we write

$$C = T_1 \Rightarrow_u T_2$$

In the case where a constraint is created or made public in a speech act u and then subsequently made public in a second speech act v , we write:

$$C = T_1 \Rightarrow_{u,v} T_2$$

Formally (i.e., beyond the notational change), this means that we have changed our conception of a constraint to include a memory of the situation in which it arose or became salient. Those familiar with situation theory will realize at once that, while this change will not affect any existing work in situation theory, it is a major change conceptually. We could have captured the necessary links by an alternative means that did not require such a fundamental change in the underlying philosophy, by (for example) incorporating the speech act that gives rise to the constraint in the prior type T_1 . But this, we felt, would be to do disservice to the very notion of speech acts we are trying to model. It is an essential aspect of the nature of a speech act that it takes place, and that should be reflected in the mathematical structure we introduce to model the act, not simply captured by a technical trick. In short, in a situation-theoretic treatment of speech acts, the origin of a constraint or the circumstances of it being made salient is just as important as its content.

With this new conception of a constraint in hand now, we modify our previous definitions for the three speech acts by replacing the old kinds of constraints by their newer counterparts. Everything else remains that same. Accordingly we will not spell out all the modified definitions here, and by way of illustration will just reconsider our first example.

This is where Alex (A) says to Betty (B):

IF GEORGE IS AT THE PARTY, GIVE HIM THIS RING.

Let u denote this speech act. By the speaker's connections, there is an individual G whom Alex refers to by the word GEORGE, a situation p referred to by the phrase THE PARTY, and an object r referred to by the phrase THIS RING. Then the constraint determined by Alex's utterance is $C = T_1 \Rightarrow_u T_2$, where:

$$\begin{aligned} T_1 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{present}, G, \dot{i}_{\text{here}}, \dot{t}_{\text{now}}, 1 \rangle\rangle] \\ T_2 &= [\dot{u} \mid \dot{u} \models \langle\langle \text{gives}, B, G, r, \dot{i}_{\text{here}}, \dot{t}_{\text{now}}, 1 \rangle\rangle] \end{aligned}$$

Exactly as before now, Alex's command is accepted if the constraint C is added to \mathcal{C}_B (by B), and is met if there is a situation v that fulfills C with respect to p . (If you

assume that Alex’s command has the implicature that Betty should give George the ring at the party, then the command will be met if p itself fulfills C with respect to p .)

Notice that the machinery we have developed above allows us to view a perlocutionary act as a perturbation of what we might call *constraint space*. The elements of constraint space are sextuples $(a, b, \mathcal{I}_a, \mathcal{C}_a, \mathcal{I}_b, \mathcal{C}_b)$, and a perlocutionary act u involving a speaker a and a listener b gives rise to a state change:

$$(a, b, \mathcal{I}_a, \mathcal{C}_a, \mathcal{I}_b, \mathcal{C}_b) \longrightarrow (a, b, \mathcal{I}_a', \mathcal{C}_a', \mathcal{I}_b', \mathcal{C}_b')$$

We shall come back to this idea later. In the meantime, having developed a mathematical framework for describing the basic communicative acts that support Internet commerce (and other activities), our next step is to analyze those acts in greater depth, with a view to understanding what it is about them that enables Internet commerce to take place. In particular — and this will be of significant relevance to intelligence analysis — how does the Internet serve as a carrier of trust? When we have completed this study, we will investigate how to make use of the results we obtain to develop further our mathematical model.

15 Web-wisdom: Evaluating trust in Internet activity

Before looking at Internet traffic in general and Internet commerce in particular (with a view to examining ways of evaluating trust), it makes sense to reflect briefly on how interactions using other forms of communication generate, sustain, or destroy trust.

Until roughly the middle of the nineteenth century, trust in a message ultimately came down to trust in either:

1. the original source of the message and the degree of trust that could be ascribed to that source (be it an individual or an organization); or
2. one or more endorsements of the message by others and the degree of trust that could be ascribed to those individuals or organizations; or
3. a combination of both of the above.

In short, trust in a message came down to trust in *the individual or organization* that either produced or endorsed (possibly by delivery) the message. And that, of course, required some knowledge of about that source or endorser. This changed with the widespread growth of, in particular, mass newspapers and, perhaps less significantly, dictionaries, encyclopedias, and reference books, during the early nineteenth century. By “commoditizing” information, delivering it in a single, recognizable format, with an institutional “stamp of authority,” these media effectively separated off the information from its source. Indeed, it was with this shift that the modern notion of *information* first began to appear, as we comment on briefly below.

According to the *Online Etymology Dictionary* (<http://www.etymonline.com/>), the word “information” goes back to the fourteenth century:

information 1387, “act of informing,” from O.Fr. *informacion*, from L. *informationem* (nom. *informatio*) “outline, concept, idea,” noun of action from *informare* (see **inform**). Meaning “knowledge communicated” is from c.1450. Short form *info* is attested from 1906. *Info-mercial* and *info-tainment* are from 1983.

Initially, then, information was thought of as attached to an individual person, something that resulted from a person being *informed*. For example, in her novel *Emma*, Jane Austin writes:

“Mr Martin, I suppose, is not a man of information beyond the line of his own business? He does not read?”

Austin lived from 1775 to 1817, and her novels are highly regarded in large part because of the deep insights into contemporary society that they display. It can safely be assumed that Austin’s use of the word “information” provides an accurate picture of the way the word was generally and widely understood at the time. The term “man of information” in the passage quoted would today be rendered as “man of learning”, or “well educated man”, or “well informed man.” While Austin would presumably have viewed the two phrases “man of information” and “well informed man” as little more than syntactic variants of each another, the meaning of “information” has changed so greatly in the intervening years that the former phrase now carries a very different connotation than in the late eighteenth and early nineteenth centuries.

With the development of new media (newspapers, dictionaries, encyclopedias, and general reference books), as well as the introduction of postal services, the telegraph, and later the telephone, it became possible to identify (or conceive of) a “substance” or “commodity” that could be called “information”.

That substance was largely autonomous, having an existence outside the individual human mind. It came in identifiable, discrete pieces. For instance, newspapers imposed the same physical structure (a block of text within a very small range of size) on every topic reported on, be it politics, war, sport, theater, science, or whatever. Moreover, the organizations that produced newspapers, reference books, and the like provided an institutional “stamp of approval” on the information they imparted, giving it the air of being neutral, free of bias and personal perspective or interpretation — “the truth,” something to be trusted.

Indeed, the degree of truth to be assigned to a report could be inferred purely from the *informational structure*, without any specific knowledge of the individual or organization that originated or delivered it. For example, Nunberg [32] points out that if you get into a rental car in a strange city and turn on the car radio and hear a sports report, you will take the report as being true — you trust that unknown reporter at that unknown station. You do so based solely on the structure of the

communication, without the need for any knowledge of the radio station, its production staff, the reporter, etc. Likewise, you are also likely to take as veridical a news broadcast heard on the same radio (although you are likely to view more skeptically an editorial commentary on a political debate). Today, we frequently, and without reflecting on the fact, endow trust to information purely on the structure — the syntax if you will — of the means by which we obtain it. Yet, this is a relatively recent phenomenon. Putting trust in a message purely on the basis of its structure in this way, with no knowledge of who produced it, was something that a person living in the eighteenth century could not have done.

One of the principal challenges facing an intelligence analyst — indeed, arguably the most significant challenge — is deciding how much trust to put in a particular piece of intelligence. As has been the case throughout history, if the source is known, and known to be trustworthy, the analyst can surely take the information to be reliable. But what if the identity of the source is unknown, or if the identity is known but it is not known if the source is trustworthy? This state of affairs is particularly common for information acquired by trawling the Internet. By allowing anyone to “be a publisher,” in the form of homepages, blogs, e-zines, and the like, desktop computers and the World Wide Web have led to a new kind of printed matter production and distribution that often has all the appearances of a regular newspaper — or at least the online version of such — without the institutional reputation to authenticate what is written. On the Web, it is no longer possible to judge the reliability of, say, a particular article, by virtue of its structure, as it was when only established newspapers had the ability to produce material with such an appearance. Today, anyone with a desktop computer and Internet access can produce and make public a document that is visually indistinguishable from the online version of the front page of the New York Times. In such an environment, the Web user has to look at the source of the article — just as was the case prior to the nineteenth century.

In the case of Web-based searches for information, how then is the intelligence analyst going to be able to locate the information he or she requires to complete an analysis, and how is he or she going to judge its reliability?

Prior to the late 1990s, even the first part of the task was prohibitively difficult. There was so much information on the Web that finding the relevant items was often all but impossible. Then along came a batch of smart search engines, the most notable and to date the most successful being Google, developed by then Stanford graduate students Larry Page and Sergey Brin in the mid 1990s. The new generation of search engines proved to be extremely successful in ranking all the items on the Web that contain the keywords specified by the user in such a way that the higher ranked items (often among thousands found for which there is a keyword match) is highly likely to be of relevance to the user. They achieved their high success rate by taking advantage of the architecture of the Web.

When a Webpage is stored on a server, along with it is stored a complete history of which users (more accurately, which other computers) have consulted that page, and when. Exactly where that traffic information is stored is not important for this

discussion. The crucial point is that the original document is transparently linked to a file that contains all that traffic information, and hence the two may be thought of as a single informational entity. Google and the other new generation search engines rank Web pages based on that traffic information.

The general idea behind present-day search engines — the exact details vary from one search engine to another, and many features are proprietary secrets — is that the importance of a particular document to a particular user is based on how many other users have consulted it. But this is not done blindly. Rather, it is done by means of a recursive procedure whereby each other user is him or herself ranked by how many other highly ranked (by this procedure) Web users consult his or her Website.

In other words, the ranking of Web pages achieved by these search engines is an accumulation of rankings *by ranked human users of the Web*. Another way to look at this, we suggest, is that, over time the Web accumulates massive amounts of what we propose to call *Web-wisdom* about the relative importance (or at least value-weighted popularity) of each Web page, and the new generation search engines capitalize on this accumulated Web-wisdom. This amounts to a modern-day Web equivalent to the old method of assessing importance by appeal to a range of knowledgeable or authoritative figures.

In the case of the relevance of information on the Web, the new search engines perform very well. Can a similar feat be performed for evaluating trust? For instance, can a way be found to mine the accumulated Web-wisdom to obtain a weighted aggregate of the trust other Web users have developed toward a particular contributor to the Web? Since trust in someone is something that is subject to development and revision based on successive interactions with that individual, any algorithm that assesses trust would presumably involve close examination of that individual's previous Internet interactions.

Since Internet interactions are entirely linguistic — indeed, the transmission of written language (symbols) — whereas trust is a human phenomenon, as a first step toward such an analysis, we should take a somewhat detailed look at the way written language is able to convey human attitudes and dispositions.

16 The symbolic transmission of attitudes and dispositions

Suppose speaker a utters a sentence Φ to b with the intention of achieving a certain outcome. Depending on that desired outcome, a will choose a particular kind of sentence Φ — an indicative in order to make a statement (i.e., to convey information to b) or an appropriate imperative in order to make a request, issue an instruction, make a request, etc.²³ On what basis does a make this choice and how does it come

²³**in-dic-a-tive** Pronunciation: in-'di-k&tiv Function: *adjective* **1** : of, relating to, or constituting a verb form or set of verb forms that represents the denoted act or state as an objective fact [the *indicative* mood]

about that *a*'s utterance of Φ achieves this goal, assuming it does? To see why these are highly non-trivial questions, notice that there is not a simple one-to-one relation between sentences and the effects speakers intend to achieve by their utterances or the effects such utterances actually have.

For example, a supervisor (female) may utter the indicative sentence

I'LL BE FOLLOWING YOUR PROGRESS CLOSELY

to a subordinate (male) for a variety of purposes. Although it is possible that the speaker intends to convey the factual information that the sentence is apparently about, it is much more likely that the speaker intends to put the listener on notice that he should take particular care in carrying out the activity implicitly referred to. In fact, it might be that the speaker has no intention of following the listener's progress at all. As a simple statement, the utterance may be false, the speaker seeking only to create in the listener the *belief* that his progress will be monitored. Moreover, it is possible that the listener realizes this fact, and nevertheless feels put on notice that his progress should meet the speaker's expectation. In fact, in making the utterance, the speaker may expect that the listener will realize he is not actually going to be monitored, and yet the speech act nevertheless still achieves the desired aim.

Given such a range of possibilities, does it follow that the actual words spoken have little to do with the effect the utterance actually achieves? Surely not. After all, there is no way that uttering the sentence

I FEEL TIRED

could achieve the same result. The sentence uttered definitely matters, but the mapping from the sentence to the desired outcome (or the unintended actual outcome) is not a simple one.

First of all, what is it about a given sentence that determines the possible uses to which it may be put, i.e., the possible outcomes that may be achieved by its utterance? It's tempting to say that the key is the "literal meaning" of the sentence, but that doesn't get you very far, since the next two questions are *What is literal meaning?* and *How is it determined?*, and all that has been achieved is to change the terminology in which the question is posed. Rather, a better answer is that what determines the possible uses to which a given sentence may be put is the range of uses to which sentences of that type are *normally* put.²⁴ By normal we mean something at a level of a socio-linguistic norm, rather than statistical clustering, but in practice the latter will surely suffice.

im-per-a-tive Pronunciation: im-'per-&-tiv Function: *adjective* Etymology: Middle English *imperatyf*, from Late Latin *imperativus*, from Latin *imperatus*, past participle of *imperare* to command
1 a : of, relating to, or constituting the grammatical mood that expresses the will to influence the behavior of another **b** : expressive of a command, entreaty, or exhortation **c** : having power to restrain, control, and direct.

²⁴This analysis is influenced by Millikan [31].

Thus, for example, a *normal* use of a sentence “ A is P ” for a definite noun phrase A and a predicate P is to express a proposition of the form $\hat{P}(\hat{A})$, where \hat{A} is the referent of the utterance of A and \hat{P} is the predicate that interprets the utterance of P .

The normal use of a sentence in an utterance is the one that a hypothetical third-party observer unacquainted with the two individuals concerned, unaware of any background circumstances, and unfamiliar with any idiomatic uses of the language, would reasonably take to be the one intended.

To get a better appreciation of the issue here, consider the four ways in which a 's utterance to b of the sentence Φ might fail to accord with the normal use.

1. a intends the normal use but b does not understand it that way.
2. b understands the utterance as according with the normal use but a does not intend it that way.
3. a does not intend the normal use and b does not understand it according to the normal use.
4. a intends the normal use and b understands it according to the normal use but the utterance does not succeed in that way.

Failure of the first kind could arise either because b simply fails to understand the utterance as a intends or else is uncooperative, refusing to accept or do what a says.

Failure of the second kind most commonly arises when a sets out to deceive or trick b , either by telling a lie and getting b to accept as factual an item of information that is false (indicative sentences), or else by giving a request or instruction to b that b cannot possibly carry out successfully (imperative sentences).

The third kind of failure can arise when the speaker sets out with an insincere intention, but also common is when the speaker intends an idiomatic or secondary meaning and the listener understands it that way. For example, if Alice says to Bob, “I’ll die if you don’t give me another slice of that pie,” Alice does not mean her words to be taken literally; rather, her intention is to express great desire for a second piece of pie, and Bob surely understands it that way. This, of course, is not a *communication* failure, rather just failure to accord to normal use of the sentence uttered.

Finally, the fourth kind of failure can arise when the speaker utters an indicative sentence that she believes expresses a true proposition and the listener believes what he is told, but unbeknownst to the speaker the proposition is false. Alternatively, the speaker may make a promise (an imperative sentence) that she intends to keep, and the listener expects the speaker to do as promised, but then circumstances arise that prevent the speaker from so doing.

Now let us ask ourselves: what determines the normal uses of sentences of a given type? Given our definition of normal use, that amounts to asking: what encourages a speaker to utter a certain kind of sentence in order to achieve a given aim? The answer: speakers seek to achieve *compliance* with the utterance — either the acceptance of a

piece of information (belief formation or confirmation) in the case of an indicative sentence or the formation or confirmation of one or more constraints (expectations or intentions to perform certain acts) in the case of imperatives.

On the face of it, this seems backwards, like defining a tool in terms of how it is normally used. How can you define tool usage before you have the tool? But this is *exactly* what you encounter when you do try to define a tool! For example, what constitutes a hammer? The only comprehensive answer is that a hammer is an object you would normally use to hammer something. But what does it mean to say something is hammered? That it is struck with an object of sufficient mass and rigidity to be effective when so used. Hammers are determined by hammering. What constitutes a hammer in one circumstance may not in another. What makes the “definition” of a hammer as a device for hammering meaningful and useful is that it is based on a normal human activity: hammering. Hammers exist because people have repeated need to hammer things. Although it is possible to use a brick or a screwdriver head to hammer, those are abnormal, and people who wish to hammer something, preferentially use purpose built hammers, if they are to hand. Of course, although hammers are most suited as a tool for hammering, that does not prevent them from being used for other purposes, say a doorstop.

Likewise, a speaker who wishes to achieve a certain end will choose words the she knows from experience are most likely to accomplish that purpose. Repeated with frequency throughout a society, this serves to establish patterns of normal usage. Given a basis of normal use, it is possible for speakers to deliberately misuse certain expressions, say by lying, but such (occasional) misuse is possible (as a device that meets with the intended success) only because it goes against the norm. For instance, if most, or even sufficiently many, utterances of indicative sentences failed to deliver the truth, people would soon stop viewing them as carriers of information. A lie succeeds (as a lie) only if the listener believes the speaker is telling the truth, and that requires that the listener has good reason to expect the truth — something that will occur only if that is normally the case with utterances of that type (and perhaps by that person). Thus, for the most part, people have strong reason to adopt normal use of language.

Note that, in most cases, the meaning of a sentence that accords with normal usage is what is often called the literal meaning. But since words get their meanings from common usage — which can change over time, leading to a change in word meaning — it would be misleading to use literal meaning as the fundamental notion here. Besides, sometimes normal use is not the same as literal meaning. An obvious example is provided by negative questions, such as “don’t you?” Taken literally, this should have exactly opposite answers to “do you?” but in fact normal usage (in the English speaking world) treats the two questions as synonymous.

17 Understanding what makes Internet commerce work

Turning now to Internet communication, suppose that Alice (*A*) sends an email message to Bob (*B*) that says

PLEASE SEND ME A COPY OF DEVLIN'S NEW BOOK.

There are various ways of describing this act. At level 0, the level of the physical world, activity within Alice's brain causes various hand and finger movements, depressing certain keys on the computer keyboard, electric currents flow, visual patterns appear on the monitor screen in front of Bob, and associated patterns of light fall on Bob's retinas, causing a certain cognitive state in Bob's mind. This level of description is appropriate for an electrical engineer designing the technologies involved.

At the next level of description, level 1, Alice generates ("utters") a well formed sentence of English which Bob receives. In Austin's [2] terminology, Alice performs a *locutionary act*.

One level further up again, Alice *issues a request* to Bob by virtue of generating an imperative sentence. This is what Austin would describe as an *illocutionary act*.

Finally, if Bob understands Alice's message and accepts it, thereby forming an intention to send her a copy of the requested book, then Alice has successfully *made a request* of Bob — in Austin's terms, she has performed a *perlocutionary act*.

Information based studies of linguistic communication, such as situation semantics, focus on locutionary acts and illocutionary acts and the relationship between them. The ultimate focus of our present study is perlocution. We shall also look at illocution, insofar as we will need to understand how illocutionary acts result in perlocution.

Notice that Alice performs her locutionary and illocutionary acts even if Bob does not receive the message — say if his email system is down. But if Bob does not receive the message (and act upon it), there is no perlocutionary act. People can perform locutionary and illocutionary acts in isolation, but for a perlocutionary act there has to be a (cooperative) recipient.

We should also remark that the nomenclature reflects linguistic action, but perlocution does not require language. Alice can make Bob shut up (a perlocutionary act) either by saying "Shut up!" (an illocutionary act) or by putting her finger to her lips (a non-linguistic, physical act).

Although we have presented locution, illocution, and perlocution as increasing levels, our earlier discussion shows that perlocution is the most fundamental of the three, and that the normal use of language is determined by the perlocutions it achieves. Expressed differently, language is a communicative tool whose development has been constrained, and whose use is constrained, by the perlocutions it facilitates and supports.

Internet commerce depends on the successful completion of perlocutionary acts, achieved by illocutionary acts. Our present investigation seeks to understand how context influences this process.

Now let's go back to Alice's email request to Bob:

PLEASE SEND ME A COPY OF DEVLIN'S NEW BOOK.

With a view to modeling this request in our new model of communication (the Devlin model), we shall describe it in terms of five factors: information (INF), desires (DES), requests (REQ), commitments (COM), and expectations (EXP).

Let S denote the action of B sending DNB to A .

Recall that in the Devlin model, we seek to express perlocutionary acts by state changes

$$(a, b, \mathcal{I}_a, \mathcal{C}_a, \mathcal{I}_b, \mathcal{C}_b) \longrightarrow (a, b, \mathcal{I}_a', \mathcal{C}_a', \mathcal{I}_b', \mathcal{C}_b')$$

We shall represent the state changes in tabular form. In the table that follows, the successive actions are numbered by integers, with each action followed by a listing (subnumbered relative to the action) of the new states that result from that action. We'll explain the final column momentarily.

| | | | MET |
|-----|--------------------------------|--|-------|
| 1 | DES _A | Initial state: A desires S | ← 8 |
| 2 | A sends a request to B | | |
| 2.1 | INF _A | A has made a request to B for S | |
| 2.2 | EXP _A | A expects a response from B | ← 5 |
| 3 | B receives A 's message | | |
| 3.1 | INF _B | A desires S | |
| 3.2 | INF _B | A has made a request to B for S | |
| 3.3 | COM _B | B will perform S | ← 6 |
| 4 | B sends response to A | | |
| 4.1 | INF _B | B has confirmed to A he will perform S | |
| 5 | A receives B 's response | | → 2.2 |
| 5.1 | INF _A | B has received A 's request | |
| 5.2 | INF _A | B knows of A 's desire for S | |
| 5.3 | INF _A | B has confirmed to A he will perform S | |
| 5.4 | EXP _A | B will perform S | ← 7 |
| 6 | B sends DNB to A | | → 3.3 |
| 6.1 | INF _B | B has performed S | |
| 6.2 | INF _B | B has met A 's request | |
| 7 | A receives DNB from B | | → 5.4 |
| 7.1 | INF _A | B has performed S | |
| 7.2 | INF _A | B has met A 's request | |
| 8 | Final state: A 's desire met | | → 1 |

In order for a transaction to complete satisfactorily, each desire, expectation, request, or commitment must be met or kept. (It is possible that steps 4 and 5 could

be omitted, although normal Internet business practice in the US is for an immediate confirmation to be sent prior to fulfillment of the request.) The final column shows how each of these individual responses occur. For example, 2.2 shows that *A* expects a response from *B* to her initial email. The entry 5 in the final column of entry 2.2 shows that this expectation is met by action 5; likewise, the entry 2.2 in the final column of entry 5 shows that this action (*A* receives *B*'s response) meets the expectation in entry 2.2. A more graphical representation might draw in complete arrows from actions to the desires (etc.) that they discharge; our notation simply shows the start and finish points of such arrows.

MORE TO COME

18 Rules of trust

In this section we begin the process of formulating rules that help us understand the way trust grows and propagates through human–human interaction, particularly interaction in the realm of Internet commerce, where many of the evolved mechanisms for the establishment of trust in face-to-face interaction are absent.

The first step is to settle on what exactly we mean by trust. There are several different — though generally connected — interpretations of the word. The phrase “*A* trusts *B*” can mean:

1. (Believability) *A* will accept as true, information provided and warranted by *B*.
2. (Predictability) *A* can reliably predict what *B* will do — that is, *B* will conform to *A*'s expectation.
3. (Reliability) *A* is prepared to take a risk in performing a certain action whose success is dependent on *B* acting a certain way.
4. (Value exchange) *A* is prepared to make an exchange with *B*.
5. (Reciprocity) *A* is willing to give something to *B* now with an expectation of later repayment.
6. (Vulnerability) *A* will risk being taken advantage of by *B* (but expects that this will not happen).

Although we have given each of the above “definitions” in an all-or-nothing fashion, in terms of *having* (total) trust, in practice there are degrees of trust, and it is the notion of degree of trust that we shall analyze.

For the purposes of making a first pass through the terrain, we shall take for our notion of trust a fairly inclusive one that incorporates all of the variants listed above.

We assume that each agent or organization A has a *composite trust index function* T_A , that for any agent or organization B that A interacts with at time t assigns a value $0 \leq T_A(B, t) \leq 1$.

The composite trust index $T_A(B, t)$ is a measure of the degree of trust that A places in B at time t , where trust is taken to include all of the variants listed above. If $T_A(B, t) = 1$, then A has complete trust in B at time t . If $T_A(B, t) = 0$, then A has complete distrust in B at time t . In general, $T_A(B, t)$ will lie somewhere between 0 and 1, indicating an assignment of trust somewhere between those two extremes.

Already a caution is in order. Use of a numerical function suggests a precision that, not only do we not yet have any idea of how to achieve, but which in fact we strongly suspect is not achievable. In further developing the mathematical aspects of this work, we will have to tread slowly and carefully in introducing (or assuming!) structural properties of the unit interval. Initially, we shall make use of the fact that $[0, 1]$ is a partially (!) ordered set having a unique maximum and a unique minimum. For instance, we will talk about the value $T_A(B, t)$ increasing or decreasing from one time instant t to another, or of the value $T_A(B_1, t)$ being greater or lesser than $T_A(B_2, t)$. We do not rule out the possibility that at some later stage in the investigation we abandon the unit interval in favor of some other measuring set. For now, however, with the above caveats in place, it seems a reasonable way to begin.

The value $T_A(B, t)$ is influenced by A 's interactions with B and by information A receives from third-party sources C that pertains to $T_C(B, t)$; the degree to which $T_A(B, t)$ is altered by this information is directly dependent on $T_A(C, t)$. The rules below spell out these relationships in more detail.

It should be stressed that our rules are not mathematical axioms (although at the back of our mind we have an idea of perhaps eventually developing a more formal mathematical model of these rules). Nor are they rules that people invariably follow, either consciously or otherwise, when engaged in, say, Internet commerce. Their status is more that of observations of commonly occurring patterns of behavior.

Rule 1. *Trust from promises.* If, at time t , B promises to A to do X , then A forms an expectation $E_A(B, X, t)$ that B will do X . That expectation will be one of the influences (constraints) that guide A 's future actions. The higher the value of $T_A(B, t)$ (i.e., the trust A has in B at the time t the promise is made), the greater will be the strength of the expectation, and thus the more likely A will be to act on it.

If B subsequently does perform X , at some time t' later than t , then A will increase his trust in B , so that $T_A(B, t')$ will be greater than $T_A(B, t)$.

If B does not subsequently perform X , then at some time t' later than t , A will decrease his trust in B , so that $T_A(B, t')$ will be less than $T_A(B, t)$; in general, absent other circumstances, A is likely to reset $T_A(B, t') \approx 0$. (\approx means approximately equal.)

By way of a general observation, we note that increasing trust values is in general a slow process, whereas one failure to keep a promise, fulfill a request, etc. can lead at once to a complete lack of trust.

Rule 2. *Trust from requests.* If, at time t , A requests of B that he do X , then A forms an expectation $E_A(B, X, t)$ that B will do X . That expectation will be one of the influences (constraints) that guide A 's future actions. The higher the value of $T_A(B, t)$ (at time t), the greater will be the strength of the expectation, and thus the more likely A will be to act on it.

If B subsequently does perform X , at some time t' later than t , then A will increase his trust in B , so that $T_A(B, t')$ will be greater than $T_A(B, t)$.

If B does not subsequently perform X , then at some time t' later than t , A will decrease his trust in B , so that $T_A(B, t')$ will be less than $T_A(B, t)$.

Rule 3. *Trust from information.* If, at time t , B informs A of X , then A will (in general) form a qualified belief that X . That qualified belief will influence A 's future actions. The greater the degree of trust A has in B (that is, the higher the value of $T_A(B, t)$), the lower will be the level of doubt A has in the veracity of X .

If, at some subsequent time t' , A comes to know that X or acquires further information from a different source that supports X , then $T_A(B, t') > T_A(B, t)$. On the other hand, if, at some subsequent time t' , A comes to know that $\neg X$ or acquires further information from a different source that supports $\neg X$, then $T_A(B, t') < T_A(B, t)$. In fact, in general, the acquisition by A at some subsequent time t' of knowledge that $\neg X$ is likely to cause him to reset $T_A(B, t') \approx 0$.

Rule 4. *Trust from hearsay.* If $T_A(C, t) = \lambda < 1$ and at time $t' > t$, B informs A that C is trustworthy, and if moreover A trusts B , then $T_A(C, t') > \lambda$. If $T_A(C, t) = \lambda > 0$ and at time $t' > t$, B informs A that C is not trustworthy, and if moreover A trusts B , then $T_A(C, t) < \lambda$.

Rule 5. *Trust from aggregation.* If A learns that C_1, \dots, C_n regard B as trustworthy, and if A has no reason to distrust C_1, \dots, C_n , then A will increase his trust in B . The greater the degree of trust A has in C_1, \dots, C_n , the greater will be this increase. Also, the greater the value of n , the greater will be the same effect.

Rule 6. *Trust from authorities.* Certain individuals and organizations C , by virtue of their position in society, are such that $T_A(C, t)$ is high for almost all A in that society. Any such C is called an *authority* in the society. If C is an authority, then for any A in the society that recognizes the authority of C , $T_A(B, t) \approx T_C(B, t)$, for any B .

Rule 7. *Trust from reputation.* Certain individuals and organizations C acquire by a time t a publicly known reputation $R_C(t)$ (which may be high or low). If C has such a reputation at time t , then for any A that is aware of C 's reputation, $T_A(C, t) \approx R_C(t)$.

Rule 8. *Trust from risk taking.* If, at time t , $T_A(B, t)$ is sufficiently high for A to take a risk based on the predicted behavior of B , and if B 's subsequent action at time t' accords with A 's prediction, then $T_A(B, t'+) > T_A(B, t)$.

That is, trust between people can result from accurate prediction of the other party's future behavior. That is to say, trust can increase if a person takes an incremental risk in the relationship. (This requires that there is a no-zero probability of loss.)

Establishment of trust in this way has an obvious evolutionary origin. For instance, if a group of our early ancestors was going to attack a much larger animal, they needed to be able to count on (i.e., predict) that each person would perform their assigned role.

MORE TO COME

APPENDIX

A desert mirage: How U.S. misjudged Iraq's arsenal

By John Diamond, USA TODAY

WASHINGTON – One year before President Bush ordered the invasion of Iraq, a U.S. spy satellite over the western Iraqi desert photographed trailer trucks lined up beside a military bunker. Canvas shrouded the trucks' cargo.

Through a system of relays, the satellite beamed digitized images to Fort Belvoir in Virginia, south of Washington. Within hours, analysts a few miles away at CIA headquarters had the pictures on high-definition computer screens. The photos would play a critical role in an assessment that now appears to have been wrong – that Iraq had stockpiled weapons of mass destruction.

The way analysts interpreted the truck convoy photographed on March 17, 2002 – and seven others like it spotted over the next two months – is perhaps the single most important example of how U.S. intelligence went astray in its assessment of Saddam Hussein's arsenal. Analysts made logical interpretations of the evidence but based their conclusions more on supposition than fact.

The eight convoys stood out from normal Iraqi military movements. They appeared to have extra security provided by Saddam's most trusted officers, and they were accompanied by what analysts identified as tankers for decontaminating people and equipment exposed to chemical agents.

But the CIA had a problem: Once-a-day snapshots from the KH-11 spy satellite didn't show where the convoys were going. "We couldn't get a destination," a top intelligence official recalled. "We tried and tried and tried. We never could figure that out."

As far as U.S. intelligence was concerned, the convoys may as well have disappeared, like a mirage, into the Iraqi desert. Nearly a year after the U.S.-led invasion of Iraq, Saddam's supposed arsenal remains a mirage.

The convoy photos, described in detail for the first time by four high-ranking intelligence officials in extensive joint interviews, were decisive in a crucial shift by U.S. intelligence: from saying Iraq might have illegal weapons to saying that Iraq definitely had them.

The assertion that Saddam had chemical and biological weapons – and the ability to use them against his neighbors and even the United States – was expressed in an Oct. 1, 2002, document called a National Intelligence Estimate. The estimate didn't trigger President Bush's determination to oust Saddam. But it weighed heavily on members of Congress as they decided to authorize force against Iraq, and it was central to Secretary of State Colin Powell's presentation to the United Nations Security Council a year ago this week.

Powell argued that Saddam had violated U.N. resolutions, agreed to after the 1991 Gulf War, requiring Iraq to disarm. But David Kay, the former head of the CIA-directed team searching for Saddam's weapons, now says that Iraq got rid of most of its banned weapons about six months after the 1991 war and that, unknown to the CIA, Iraq's weapons research was in disarray over the past four years.

The failure to find biological or chemical weapons in Iraq has undercut the Bush administration's main justification for invading Iraq. And it has raised concerns that the United States is conducting a policy of pre-empting foreign threats with an intelligence system that is fundamentally flawed.

An independent commission, reluctantly backed by the Bush administration, will be established to find out what went wrong. Such a panel is sure to explore whether, like thirsty travelers seeking an oasis, the U.S. analysts were looking so hard for evidence of banned Iraqi weapons that they "saw" things that turned out to be illusions.

Major findings

How could the nation's \$40 billion-a-year intelligence apparatus, focused on Saddam's regime for more than a decade, have been so wrong? A three-month examination by USA TODAY of prewar intelligence on Iraq, involving more than 50 interviews and examination of thousands of pages of documents, found that:

- Volumes of intelligence suggested illegal Iraqi weapons activity but did not prove Iraq had such weapons. The evidence was intriguing but inconclusive. Spy satellites photographed convoys but couldn't determine where they were going. Human sources told of Iraqi attempts to buy banned equipment but didn't say whether the deals went through. Electronic intercepts exposed Iraqi concealment but didn't explain what was being hidden.
- Despite the lack of proof, CIA Director George Tenet and his top advisers decided to reach a definitive finding. Based on experience with Iraq – and with the Sept. 11, 2001, terrorist attacks in mind they were far more worried about underestimating the Iraqi threat than overestimating it.
- Few officials in U.S. intelligence, Congress or the executive branch seriously considered Iraq's claim that it had gotten rid of its weapons. Scarcity of evidence, intelligence officials said, stemmed not from innocence but from Iraqi concealment and lies.

The five men who put together the October 2002 intelligence estimate insist that the White House didn't pressure them into elevating the assessment of the Iraqi threat. But they were haunted by past failures and the fear of the worst-case scenario. Tenet, who declined to be interviewed for this article, pushed them to avoid wishy-washy conclusions. And they were aware that any finding exonerating Iraq would put them into conflict with top administration officials.

Now these analysts face another kind of worst-case scenario in which a war was premised on faulty analysis and their judgments are no longer trusted.

Burned before

U.S. intelligence analysts were reluctant to give Iraq the benefit of the doubt because Saddam had fooled them before.

After the 1991 war, U.N. weapons inspectors, tipped off by an Iraqi defector, uncovered a much more extensive program to develop nuclear weapons than the CIA had estimated. It

happened again in 1995 when Iraq admitted to a biological weapons program undetected by U.S. intelligence.

“The lesson of ’91 was that (Saddam) was much more effective at denial and deception than we understood, and consequently he was a lot further along than we understood,” Stuart Cohen, vice chairman of the National Intelligence Council, a senior advisory board, said in an interview.

Virtually all of the CIA’s recent, painful lessons revolved around the failure to detect and warn of a threat. These included a bombing at the Khobar Towers military barracks in Saudi Arabia in 1996; nuclear tests by India and Pakistan in 1998; the bombing of the USS Cole in Yemen in 2000; and, most traumatically, the Sept. 11 attacks.

In July 1998, a commission led by Donald Rumsfeld, who would become Bush’s Defense secretary, cautioned that U.S. intelligence might not be able to warn of emerging ballistic-missile threats from states such as North Korea, Iran and Iraq. The solution, the panel advised, was a new kind of analysis to “extrapolate a program’s scope, scale, pace and direction beyond what the hard evidence at hand unequivocally supports.”

As Defense secretary, Rumsfeld would insist that war in Iraq was waged on solid intelligence. Increasingly, however, it appears that U.S. intelligence followed the course set by Rumsfeld’s 1998 panel in extrapolating the scope of the Iraqi threat “beyond . . . the hard evidence at hand.”

Decisive convoy photos

Of all the Bush administration accusations about Iraq, none was more important than the charge that Saddam possessed chemical and biological weapons capable of killing millions of people. And no evidence was more important to making that charge than the convoy photographs taken in March, April and May 2002.

The story of the suspicious convoys in the Iraqi desert illustrates how the CIA turned tantalizing evidence of Iraqi weapons into conclusions that went beyond the available facts. It also underscores the limits of technical intelligence. Orbiting U.S. spy satellites provide periodic snapshots but, because they don’t hover over a spot on Earth, they can’t send back motion pictures of what’s going on.

The eight suspicious convoys bore a striking resemblance to known chemical-weapons convoys that had been picked up by spy satellite photos in 1988. Briefing top officials at CIA headquarters, analysts placed examples of the old and new photos side by side on poster board. They also contrasted the eight suspicious convoys with more than 100 conventional Iraqi military shipments also photographed during the spring of 2002. They showed them on posters labeled “Normal Activity” and “Unusual Activity.”

“There’s some stunningly good evidence about what I would call chemical weapons munition trans-shipment activity,” said Cohen, who played a key role in producing the Iraq intelligence estimate. Cohen said the evidence “was certainly subject to alternative interpretations, but there were very sensitive signatures involved that would have led any reasonable person to the same conclusion that we came up with.”

Another high-ranking intelligence official called the convoy images “an extraordinarily important piece. It’s one of those ‘dots’ without which we could not have reached that judgment that Saddam had restarted chemical weapons production.”

By September, after intense debate, opinion solidified, and senior analysts preparing the intelligence estimate judged with “high confidence” that the convoys carried chemical weapons. Their conclusion was timely because Bush was just then ratcheting up his case against Iraq to

the U.N. and Congress. Between October 2002 and the U.S.-led invasion the following March, satellite images showed three more convoys bearing what appeared to be the special signatures of chemical weapons.

Weeks before the invasion, however, there were signs that the CIA might be mistaken. U.N. inspectors visited the sites where the convoy photos were taken and scores of other locations, but they found no trace of chemical or biological weapons. At the CIA's prodding, the inspectors looked for decontamination trucks but reported finding standard water tankers with no evidence of decontamination gear.

Since the war, no decontamination vehicle has been found, the four intelligence officials said. U.S. interrogators have questioned scores of Iraqi military truck drivers. They either say they know nothing or tell stories that don't check out, according to a Pentagon official with knowledge of the search effort.

Deductive reasoning

What were the convoys doing if they weren't moving chemical weapons? The tanker trucks might have been carrying water in case munitions exploded, or fuel to keep a long-distance convoy moving. The trailer trucks might have been loaded with conventional rockets or shells, which would be hard to distinguish from chemical munitions. U.S. intelligence did not know for sure, and still does not know, where the convoys were going or what they were carrying.

Other critical parts of the case against Iraq were also based on deductive reasoning.

Once Iraq showed it knew how to make chemical weapons in the 1980s, U.S. intelligence assumed it held on to the recipe. "Iraq's knowledge base is absolutely critical," Cohen said. "Knowledge is not something you lose."

Beginning in 1999, spy satellite photographs taken based on tips by human sources showed that Iraq was expanding a chemical plant near Fallujah called Habaniyya II that could produce phenol and chlorine, ingredients for chemical weapons. The CIA had information from 15 people over four years saying that Iraq was reviving its weapons production capability at Habaniyya and other plants. But the CIA rated the five best of those sources as having only "moderate reliability."

Electronic intercepts and reports from human sources showed that senior officers at some of these facilities were the same people known to have been involved in Iraqi chemical weapons production in the 1980s.

Chlorine can be used for civilian purposes such as water purification. But CIA analysts remained suspicious because of reports that Iraq had a surplus of chlorine at its water treatment plants. Why expand a chlorine plant if there was a surplus, they asked, unless it was to make weapons?

The CIA detected efforts by shadowy middlemen, negotiating with foreign governments and businesses, to buy equipment and chemicals useful in making weapons of mass destruction. Without hard evidence, U.S. intelligence decided it had to assume that some illegal material was getting through, the four high-ranking intelligence officials said.

Analysts made similar assumptions from U.N. reports. U.N. inspectors, for example, said Iraq could not account for about 3,000 tons of chemicals that could be used to make weapons. CIA weapons experts said Iraq could use those chemicals to make 100 to 500 tons of chemical agent, a figure used repeatedly by administration officials. The U.N. also said that Iraq had failed to account for growth media sufficient to make up to 25,000 liters of the biological agent anthrax and that there was a "strong presumption" that 10,000 liters of anthrax Iraq had in 1991 still existed.

U.S. intelligence merged debatable intelligence about chemical and biological agents with equally debatable intelligence about weapons delivery systems. Iraq, the CIA said, still had 20 Scud missiles and was developing drone aircraft that might be launched, possibly off a merchant ship, to strike the United States.

Bush administration officials then translated the CIA's worst-case calculations into potential mass casualties. In his 2003 State of the Union address, Bush cited the U.N. figures in saying that the anthrax would be enough "to kill several million people" and that the chemical weapons could "kill untold thousands."

Powell, in his presentation last year to the U.N. Security Council, said even a conservative estimate would give Saddam enough chemical agent to attack "an area five times the size of Manhattan."

Since the war, no Scud missiles have been found. The drone aircraft U.S. search teams have found in Iraq were too small to deliver chemical or biological weapons.

'Mountain' of evidence

It is only beginning to become clear that information about Iraqi weapons was scarce because the weapons didn't exist. Aris Pappas, a former CIA analyst, said in an interview that U.S. intelligence had essentially "gone blind for three years" in Iraq after U.N. inspectors left at the end of 1998. Based on the available evidence, analysts probably made sound judgments, said Pappas, a member of an Iraq intelligence review panel established by Tenet. But they overlooked alternative explanations and paid too little heed to the weakness of their raw data.

"They keep referring to a 'mountain' of evidence. . . . But it was corroborative evidence," Pappas said, meaning evidence that supported allegations of an illegal arsenal without proving its existence.

The Bush and Clinton administrations, foreign intelligence services, and Republicans and Democrats in Congress all took it as a given that Iraq had chemical and biological weapons.

"If we were massively wrong," said Robert Einhorn, who worked on proliferation issues at the State Department in the Clinton and Bush administrations, "we were all massively wrong. Everybody."

Bush didn't believe that U.N. inspectors had forced Iraq to get rid of its banned weapons after the 1991 war. Indeed, Bush's policy assumed that U.N. inspections couldn't work. After the Sept. 11 attacks, the watchwords at the White House and CIA headquarters were, assume the worst.

"We put the analysts under tremendous pressure," said Kay, the former head of the post-war weapons search. "There is a point where an analyst simply needs to tell people: 'I can't draw a conclusion. I don't have enough data. Go get me more data.' But in the wake of 9/11, believe me, that is difficult to do."

Sept. 11 showed the consequences of failing to warn of an imminent threat. Now U.S. intelligence is grappling with the consequences of perceiving a threat that was not there.

References

- [1] Attardi, G. and Simi, M. Proofs in context, in *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth Conference* (1994).
- [2] Austin, J.L. *How To Do Things With Words*, Clarendon Press (2nd edition, 1975)

- [3] Barwise, J. *The Situation in Logic*, CSLI Lecture Notes 17 (1989).
- [4] Barwise, J. & Perry, J. *Situations and Attitudes*, MIT Press (1983).
- [5] Barwise, J. & Etchemendy, J. Information, Infons, and Inference, in Cooper, Mukai, and Perry (editors), *Situation Theory and It's Applications, Vol. 1*, Stanford University: CSLI Lecture Notes 22, 1990.
- [6] Barwise, J. & Etchemendy, J. *Hyperproof*, CSLI Publications, 1994.
- [7] Barwise, J. & Etchemendy, J. *Heterogeneous Logic*, in Glasgow, J., Narayanan, H, & Chandrasekaran, B. (1995). *Diagrammatic Reasoning: Cognitive and Computational Perspectives*, AAAI Press and MIT Press, (1995), 211-234.
- [8] Barwise, J. & Etchemendy, J. *A Computational Architecture for Heterogeneous Reasoning*, U.S. Patent 5,999,182.
- [9] Barwise, J. & Seligman, J. *Information Flow: The Logic of Distributed Systems*, Cambridge University Press (1997).
- [10] Buvac, S and Mason, I. Propositional logic of context, in *Proceedings of the Eleventh National Conference on Artificial Intelligence, Washington, D.C.* (1993).
- [11] Clark, H. *Arenas of Language Use*, University of Chicago Press (1993).
- [12] Clark, H. and Carlson, T. Context for comprehension, in Long, J. and Baddeley, A. (eds) *Attention and Performance IX*, Erlbaum (1981), pp.313–330.
- [13] Devlin, K. *Logic and Information*, Cambridge University Press (1991).
- [14] Devlin, K. *Goodbye Descartes: The End of Logic and the Search for a New Cosmology of the Mind*, John Wiley (1997).
- [15] Devlin, K. *InfoSense: Turning Information into Knowledge*, W. H. Freeman (1999).
- [16] Devlin, K. *A framework for modeling evidence-based, context-influenced reasoning*, presented at CONTEXT '03, Stanford University, June 2003, to appear.
- [17] Devlin, K. & Rosenberg, D. *Language at Work: Analyzing Communication Breakdown in the Workplace to Inform Systems Design*. Stanford University: CSLI Publications and Cambridge University Press (1996)
- [18] Farquhar, A. and Buvac, S. Putting Context Logic into Practice, technical report of the Stanford Knowledge Systems Laboratory, January, 1997.
- [19] Farquhar, A, Dappert, A, Fikes, A. and Pratt, W. Integrating Information Sources Using Context Logic, technical report of the Stanford Knowledge Systems Laboratory, January 1995.
- [20] Garfinkel, H. *Studies in Ethnomethodology*, Prentice–Hall (1967).

- [21] George, A. *Presidential Decisionmaking in Foreign Policy: The Effective Use of Information and Advice*, Boulder, CO: Westview Press (1980).
- [22] Giunchiglia, F. Contextual reasoning, *Epistemologia*, XVI (1993), pp.345–364.
- [23] Guha, R.V. *Contexts: A Formalization and Some Applications*, Ph.D. thesis, Computer Science Department, Stanford University (1991).
- [24] Gumpertz, J. & Hymes, D. (eds). *Directions in Sociolinguistics, The Ethnography of Communication*, Holt, Rinehart and Winston Inc. (1972).
- [25] Harris, W. *Interpretive acts: in search of meaning*, Clarendon (1988).
- [26] Heuer, R. J, Jr. *Psychology of Intelligence Analysis* re-published by the Central Intelligence Agency in 1999, currently available only on the Web, at: <http://www.cia.gov/csi/books/19104/>
- [27] Models of interaction of language and social life, in Gumperz J. and Hymes, D. [24] pp.35–71.
- [28] Johnston, R. *Analytic Culture in the U.S Intelligence Community: An Ethnographic Study*, CIA (2005).
- [29] McCarthy, J. Generality in artificial intelligence, *Communications of the ACM*, 30 (12), 1987, pp.1010–1035.
- [30] McCarthy, J. Notes on formalizing context, in *Proceedings of the Thirteenth International Joint Conference in Artificial Intelligence*, Chamberry (1993).
- [31] Millikan, R. *Language, Thought, and Other Biological Categories*, MIT Press (1984).
- [32] Nunberg, G. (ed), *The Future of the Book*, University of California Press (1996).
- [33] Piatelli-Palmerini, M. *Inevitable Illusions: How Mistakes of Human Reason Rule Our Minds*, John Wiley (1996).
- [34] Sacks, H. On the Analyzability of Stories by Children, in Gumperz J. and Hymes, D. [24], pp.325–345.
- [35] Searle, J. *Speech Acts*, Cambridge University Press (1969).
- [36] Shoham, Y. Varieties of context, in Lifschitz, V. (ed) *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, Academic Press (1991), pp.393–408.
- [37] Sowa, J. Laws, Facts, and Contexts: Foundations for Multimodal Reasoning, <http://www.jfsowa.com/pubs/laws.htm>
- [38] Thomason, R. Type Theoretic Foundations for Context, Part 1: Contexts as Complex Type-Theoretic Objects (1999), preprint available for download at <http://www.eecs.umich.edu/~rthomaso/documents/context/>

- [39] Thomason, R. Contextual Intensional Logic: Type-Theoretic and Dynamic Considerations (2001), preprint available for download at <http://www.eecs.umich.edu/rthomaso/documents/context/>
- [40] Yamada, T. An Ascription-Based Theory of Illocutionary Acts, in Vanderveken, D. & Kubo, S. (eds) *Essays in Speech Act Theory*, Pragmatics and Beyond, New Series 77, John Benjamins (2002), pp.151–174.
- [41] Vaughan, D. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*, University of Chicago Press (1996).

Keith Devlin
CSLI, Stanford University
Stanford, CA 94305.