Conversational Alignment

Austin Henderson
Scalable Conversations | henderson@rivcons.com

Jed Harris
Scalable Conversations | jed@jive.com

People invent and revise their conversation midsentence. People assume they understand enough to converse and then simply jump in; all the while they monitor and correct when things appear to go astray from the purposes at hand. This article explores how this adaptive regime works, and how it meshes with less adaptive regimes of machines and systems.

A Tale of Two Stories
A colleague told us a story of two friends discussing euthanasia. At least, that was what one thought they were discussing. The other heard the discussion as being about "youth in Asia." Remarkably, the conversation went on for more than five minutes before the misalignment was detected.

The "Who's on first?" comedy routine by Abbott and Costello is based on a similar misalignment. Those master comedians make the audience a knowing third party to the difficulties.

The usual accounts of such conversations would have it that this is an exceptional case, and usually speakers are well aligned. These accounts hold that good (or even perfect) alignment is necessary for conversation.

We explore an alternative perspective: These stories of mis-aligned conversations are not different in kind from more typical, apparently well-aligned conversations. Rather, we hold that all interactions are necessarily misaligned to some degree, and the mechanisms that make conversation "good" are not those that bring speakers into perfect alignment, but rather those that maintain a degree of alignment appropriate for the situation. The work of being a good conversant is to produce alignment that is just good enough for the purposes at hand.

Getting Started
If you were starting a conversation with a Martian, you might reasonably be uncertain about what you could assume concerning the Martian's view of the impending conversation—its views on inter-actional moves, language, subject matter, even what a conversation is. You would have difficulty knowing where to start.

In contrast, when you meet a colleague in the hallway, you usually get started with little difficulty. You assume they will speak, using the same language you used yesterday when you two last spoke; that a friendly greeting is a good starting subject matter; and that the conversation will be composed of both of you taking turns, some-times overlapping, with an end in the not too distant future.

We argue that the starting situations with your colleague and with the Martian are different only in degree, not in kind. In each case, both of you make a set of assumptions about the situation. And then one or other (or both!) of you will simply make some interactional move. The Martian might wave its ears; your friend might say, "Did you have a good weekend?" And as a result of that first move of plunging in, you immediately have all sorts of information that you can use as evidence for or against the assumptions you made about the conversation. Yes, the conversation appears to be talking (rather than ear waving, or crying, or hugging, or...); yes, it appears to be in English (although no doubt you may have on occasion started a conversation with "Bonjour!" to a friend who you know also speaks French); yes, it appears to be starting with social niceties; and yes, we seem to be embarking on a hallway conversation.

There is nothing determined about any of this. The world we live in emerges as we live it, and we have to take it as it comes, and make of it what we can. So you have to start with assumptions, engage in conversation on the basis of those guesses, and subsequently adjust
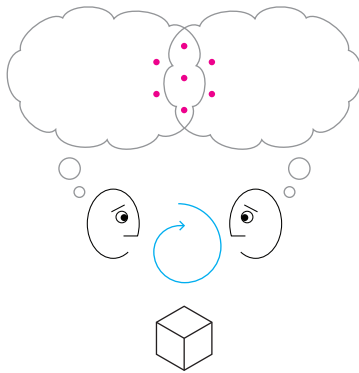
▶ We describe conversation *as if* we are creating a shared mental model...
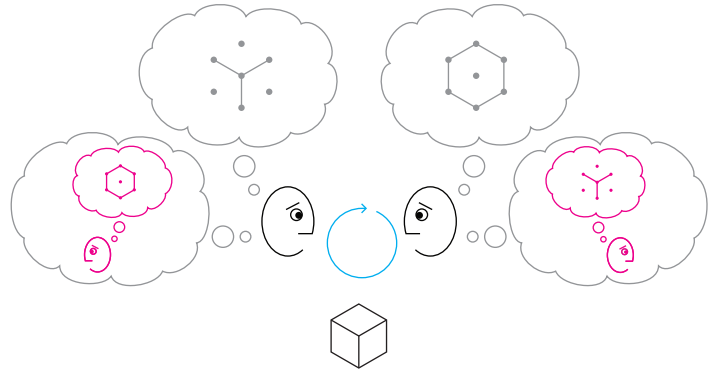
▶ But 'sharing' is shorthand for a more complex process in which we form a mental model of our interlocutor's mental model of the topic at hand,...

your assumptions as you produce evidence from the engagement.

And at the same time, your partner in this game is doing exactly the same thing: starting with assumptions, engaging, and using your conversational moves as evidence for adjusting those assumptions.

Adequate Alignment

As the conversation continues, both of you make conversational moves and monitor each other to see if you make sense out of each other's moves. In the normal (normative) case, the moves provide evidence that supports, extends, or incrementally changes the assumptions with which you started.

At the same time, both of you are monitoring each other to see whether you are being "understood"— whether the other person appears to be making enough sense out of what you said. You cannot read their mind. However, their responses are evidence of whatever sense they made of your move.

In a similar vein, when you are listening, in order to provide information for your conversational partner's use, you may signal that you are making sense of their moves: Maybe you make eye contact, give a nod or a smile, even engage in an overlapping completion of their sentence.

We achieve continued conversation by maintaining mutual assur-

ance that each of us can make enough sense of each other's moves.

Trouble

However, sometimes making sense is not so easy. In the "Who's on first?" routine, the evidence of trouble is immediate and profound. In the "euthanasia" scenario, trouble took surprisingly long to emerge.

Confronted with trouble, the next conversational move may address not whatever is under discussion, but rather the difficulty in interacting. This may take the form of "What are you talking about?" or a furrowed brow, or a conversational turn about the trouble: "When you say 'euthanasia,' are you talking about assisted dying?"

Conversation analysts refer to such shifts in subject matter from the matter at hand to the conversation itself as "breakdowns." A breakdown in this sense is a response to a feeling that our interaction is not working well enough, and that the conversation should be interrupted and refocused on the conversation itself. When a hammer handle breaks, fixing the roof stops, and fixing the hammer begins. We shift focus to converse about the conversation and "repair" the breakdown.

Once a repair has been concluded, the conversation can pick up where it left off, but now possibly with improved alignment—a better grip on the mechanics of convers-

ing, the meaning of the terms, even the purpose of the discussion.

Levels

In conversation we always work on multiple levels: We monitor the comfort, interest, and comprehension of our partners; adjust our approach; maybe switch topics, etc. Explicit repair of breakdowns is an unusually clear case of switching primary attention to a different level.

In most situations we shift our emphasis between these levels so easily that we are hardly aware they exist and so we may find it difficult to make our multilevel negotiations explicit.

Sense Making

We are all very good at making sense of situations, fitting things into the context and moving along. The sense we have made may later turn out to be flawed, but we are really troubled only if we can't make our understanding work well enough for the purposes at hand.

Sometimes we find that new activity is confirming evidence: We can make sense of it without any change to our assumptions or understanding. It fits right into the sense we have made of the world.

Alternatively, we may have to change our assumptions in order to make sense of a move. We might think the sky is blue, and our
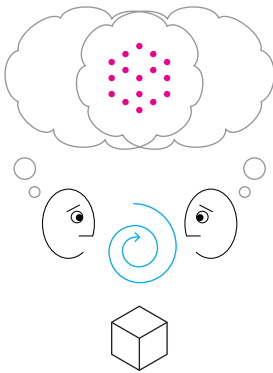
▶ *...as if* our mental models are converging.
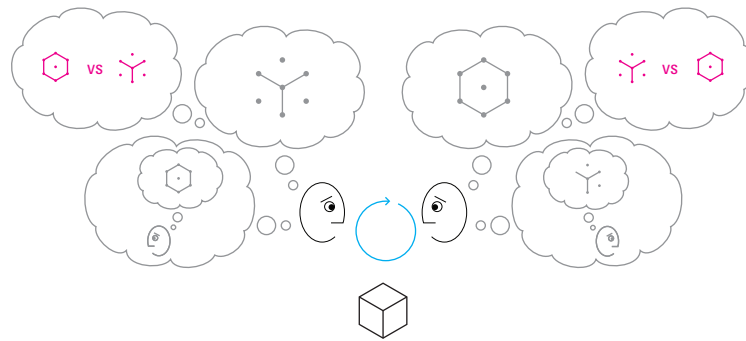
▶ ...and compare it to our own model in order to confirm sufficient agreement on relevant areas to continue the conversation.

conversational partner might say, "Looks like rain." On observation, low clouds in the west are indeed there, so we adjust our blue sky to have low western clouds, and we adjust our assumptions about our partner to reflect that they see the world that way too.

When something doesn't fit, we tend to look for the smallest (and often most local) changes in our view that will have things make sense. After which, we may opt to move on. But we also often retain a concurrent view of how well we are doing in making sense of things, how much work we had to do, how happy we were with the result, and whether there are loose ends—simply, is the conversation working?

Because it is expensive to drastically reset our assumptions, we are inclined to delay doing so until we are reasonably sure about being unsure. Therefore, the suspicion of misalignment often develops over a number of interactional moves, finally reaching the point where we feel the effort of realignment is worthwhile.

This process of working within common assumptions, noting anomalies, seeking the smallest changes that can get us back on a track that seems to make sense, and sometimes reluctantly accepting the need for a more radical overhaul of our conceptual framework exactly fits the pattern

Thomas Kuhn first described in *The Structure of Scientific Revolutions*—though on a much smaller scale. Each partner has their own tacit, informal theory of the conversational ground, and the interaction proceeds by growing and/or challenging the partners' theories. In our design conversations, our local renegotiation of meanings often ripples out to shift our design goals, directions, and fantasies, and in the most fruitful cases may pave the way to revolutions.

Repair
When we shift focus to improve alignment, we are working to repair the breakdown:

A; "*Bonjour!*" (start shift)

B; "Oh, *parlez-vous français?*" (start repair)

A; "No, but I grew up in Toronto and struggled with French for five years in high school."

B; "Oh, I see. (end repair) OK. (end shift) *Bonjour* to you too."

And, of course, shifts and repairs are themselves conversation. You and your conversational partner have to deal with them in exactly the same way as any other conversation—including the ones in which you encountered a breakdown. You have to use the same conversational mechanisms and practices. In tough cases, when implicit coordination breaks down, you have to hope your partner recognizes that you are

shifting focus and talking about the talk, not about the weather. You have to make assumptions, monitor, adjust, and continue. You have to work to stay adequately aligned through this sub-conversation and to get back to the interrupted one.

Uncertainty
In talking about the conversation, you are using the same assume-act-monitor-adjust style of communicating as in any other conversation. And you get only circumstantial evidence that you are understanding what sense your partner is making of the whole thing.

When you work on terminology and meaning and philosophical frameworks, you may infer a lot about the alignment of your respective views. However, you cannot ever know for sure what sense your partner is making, nor how closely aligned that sense is to the sense you are making.

Aligned Enough
Fortunately, you don't need to know your partner's sense of the conversation precisely or certainly. You need only enough evidence to stay confident that your alignment can meet the needs of the conversation. Small talk about having a nice day will probably not require exploration of a partner's sense of the terms of meteorology. But discussion of a hurricane might.

INFORMATION THEORY

The work that defined information theory is Claude Shannon, A mathematical theory of communication, *Bell System Technical Journal 27* (1948), 379–423, 623–656. Since then, of course, it has become the foundation of entire disciplines.

COGNITIVE LINGUISTICS

Deirdre Wilson and Dan Sperber, Relevance theory, in *The Handbook of Pragmatics*, L.R. Horn and G. Ward (eds.), Wiley-Blackwell, Oxford, UK, 2004, 607–632; http://www.dan.sperber.fr/?p=93/.

EVOLUTIONARY LINGUISTICS

Luc Steels, Grounding symbols through evolutionary language games, in *Simulating the Evolution of Language*, A. Cangelosi and D. Parisi (eds.), Springer, New York, 2001.

GAME THEORY

Computational approximation of multi-agent game equilibria, such as Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*, Cambridge University Press, Cambridge, UK, 2009; http://www.masfoundations.org/download.html./

H. Peyton Y*oung, Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, Princeton University Press, Princeton, NJ, 1998. Analyzes convergence on consensus using game theory.

BREAKDOWNS

Analysis of breakdowns in interaction has many roots, including work by Heidegger, Garfinkle, Schegloff, Bodker, and others. For a review of "trouble and repair" in this context, see B. Jordan and A. Henderson, Interaction analysis: Foundations and practice, *The Journal of the Learning Sciences 4*, 1 (1995), 39-103.

SCIENTIFIC REVOLUTIONS

Kuhn's work catalyzed an entire field devoted to these issues. His original statement of the pattern was Thomas S. Kuhn, *The Structure of Scientific Revolutions,* University of Chicago Press, Chicago, 1962.

ADAPTIVE COMMUNICATIONS

This area is under very active development. The first comprehensive treatment was Peter D. Grünwald, *The Minimum Description Length Principle,* MIT Press, Cambridge, MA, 2007. However, research has progressed significantly beyond Grünwald's account.

Your understanding of the purpose of the conversation will tell you how much alignment is needed and how hard you need to work at achieving it. And of course your partner will have their own view of the conversational purpose and their willingness to invest in achieving alignment. Their view may be different. How different? Recursively, the answer is: however much each of you find sufficient for the purposes at hand.

Stability

As the conversation continues, confidence in sufficient alignment can build and be reinforced by the success of the preceding talk: The same term continues to be used in ways that continue to make the same sense; conversational moves do not lead to incompatible responses, and any breakdowns are easy to repair. Overall, a feeling of stable convergence can develop.

We may think of this as a "fixed point" of the conversational negotiational activity, in the mathematical sense that the ongoing conversation keeps converging on the same underlying understanding while continuing to add layers and details to that understanding.

Further, stability can accumulate. Each discussion means the assumptions for starting the next discussion can be better, convergence can be faster, and so forth. This is sometimes referred to as "having good bandwidth" with someone. Indeed, if our communication channel is fixed—for example, face-to-face conversation—we get greater effective bandwidth. Conversely, if we just want to convey a specific point, we can do it with less bandwidth. This metric has been partially formalized in some three-level accounts of adaptive communication.

As the background becomes stable, we are increasingly tempted to treat it as if it were frozen forever. This can make it difficult for us to "challenge the brief," to question and revise the context of our own designs. Great designs typically involve unfreezing and renegotiation of the background.

Codes and Negotiations

Fixed points in conversation remind us of classical information theory, which starts from the premise that communication always depends on a fixed "code" that defines the possible messages and the encoding of those messages in the channel. Information theory was inspired by the experience of building a national telephone network and has subsequently become the standard basis for designing machine-machine interactions.

In our view, this is an optimized case of collapsed negotiation-based conversation, with completely stable fixed points of conversational meaning. This raises two questions for us: Where did the codes come from, and how can codes change?

Where did the code come from? Information theory is concerned with optimizing communication efficiency in a static environment. As mentioned earlier, in conversations based on stable understandings, fixed points—the codes—can be frozen and sedimented.

How can codes change? In code-based communication there is no place for negotiation of the codes, so system-builders must negotiate outside the code itself to respond to misalignment. Such negotiation mechanisms need to be included in a full account of how codes work in the real world. That is where our "larger" perspective is required.

Consider HTML. A given version of HTML may be viewed as a classi-

cal information theoretic code, but in practice HTML is defined by an ecology of roughly compatible codes being generated and accepted by multiple (buggy) software packages, and furthermore constantly being renegotiated at higher levels by developers, standards bodies, and so forth. We have to consider multiple levels to understand the evolution or even the current status of HTML.

While our view is unusual in most parts of computer science, powerful conceptual tools are available to support it. It has been explored in different forms in cognitive linguistics and has been formally analyzed in various ways using game theory.

So both perspectives are necessary; they complete each other. Negotiated systems can gain efficiency from stability when it has emerged, and code-based systems need negotiation, so that they can be responsive to a diverse and changing world.

Change
Because conversation does not depend on preestablished agreements, and the mechanisms of monitoring and repair help us handle a partner's conversational moves that we can't understand, this conversational practice is also suitable for dealing with a changing world.

If a partner changes their mind about something—and that change is relevant to a discussion—the mechanisms for conversation have the capacity for detecting the mismatch from the conversational moves, shifting focus, negotiating adequate realignment, and resuming.

Agreement
We often say we "reach agreement" with others on some matter. We talk as if there is a view that we then all share (a "common ground"). In contrast, our view is that the idea of "reaching a shared view" is a linguistic gloss, shorthand for something much more complex and powerful. Agreement is not a single ground. Rather, it is a commitment to continue to work together to maintain coherence.

We would say the parties to an agreement interact with each other until they each can construct senses for themselves and for each other that are aligned enough, so they anticipate their subsequent individual actions will be coherent enough to achieve their goals.

A common failing of meetings is that participants engage in "collaborative misalignment"—working hard to get language that all can agree to but avoiding testing whether the inevitably disparate senses carried away will lead to collectively coherent action. Another failing is that on later encountering a world that was unanticipated during the meeting, individual action is based on personal understanding alone rather than on the personally aggregated sense of the disparate understandings of all.

Coherence, Responsiveness, and Scale
Finally, we see this perspective as strongly supporting the need for systems to both be responsive to many particular viewpoints and also to achieve coherence in activity, and to do so even as scale increases.

Consider scaling the achievement of conversational alignment over many people doing many things. Meanings, purposes, and negotiations are local, but because of overlapping alignments, they begin to cohere into a commonality that we think of as the meaning of language—again, at risk of reverting to the one level code perspective. We believe it is important to stay aware that this sense of commonality is a gloss for a vast dynamic network of local exchange and negotiation of meaning. Our systems must support both the efficient use of commonality and the renegotiation of meaning when the commonality is inadequate to the needs of participants.

Design
Unlike communications systems, people interact with each other without first agreeing on communication protocols. This is possible because they start and continue to act on the assumption that they understand enough to communicate—and then they interact. All the while they monitor and correct when things appear not to be working well enough for the purposes at hand.

As designers, conversations are at the center of our practice. Now we must challenge ourselves to design systems that accept and support users' conversations. Machines cannot yet negotiate alignment, but they can and should help their users carry on conversations, recognize breakdowns, and negotiate meanings to meet the needs of a heterogeneous and changing world.

ABOUT THE AUTHORS
Austin Henderson's 45-year career in HCI includes research, design, architecture, product development, and consulting at Lincoln Laboratory, BBN, Xerox (PARC and EuroPARC), Fitch, Apple, and Pitney Bowes. He focuses on technology in conversations in a rich and changing world.

Jed Harris started out exploring cultural anthropology, linguistics, philosophy of science, and artificial intelligence research, and then spent 40 years in research and development at SRI, Stanford, Xerox PARC, Data General, Intel, and Apple. He's now happily meshing technology with the human sciences.