

The BIC Model: A Blueprint for the Communicator

J.P. de Ruiter and N.J. Enfield

Max Planck Institute for Psycholinguistics

Abstract. In this paper, we outline a cognitive architecture for communicators, called the BIC model. The model consists of three main components. First, a (B)iological component, in which the genetic or built-in capacities of the communicator are specified. Second, an (I)nteraction Engine which uses neo-Gricean mutual simulation to attribute communicative intentions to signals, and to create signals to convey communicative intentions to the I-system of other agents. The third component of the BIC model is a content addressable database of (C)onventions which is used to store form/meaning mappings that have been successfully computed by the I-system. These stored form/meaning mappings are indexed by types of communicative context, so they can be retrieved by the I-system to save computational resources. The model can be used both as a computational architecture for a communication module in an artificial agent and as a conceptual model of the human communicator.

1 Introduction

Claude Shannon's [1] influential treatment "A mathematical model of communication" still influences the way many cognitive scientists think about communication. However, Shannon's theory is a theory of *data transmission*, not of real communication between two autonomous agents. The point of his work was to show how a series of bits could reliably be transmitted over a channel of limited capacity, and how this data transmission could be made less susceptible to the main enemy of communication engineers: noise (see Figure 1).

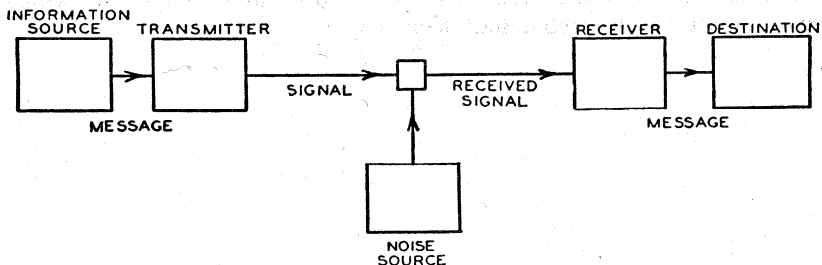


Fig. 1. Shannon's (1948) famous diagram

When applied to communication between agents (either human or artificial) the transformations that take place at the transmitter and receiver's end take on a

Saussurean character: for successful communication, there has to be an unambiguous mapping of signals (e.g., bit streams or speech sounds) to meanings, and a mapping of meanings onto signals that is its inverse. In psycholinguistic models of communication, this assumption is reflected at the level of phonology, syntax, and semantics (see e.g. [2]). In artificial communicators this assumption still guides implementations of speech recognition, comprehension, and generation. These levels of representation enable the communicator to unambiguously map signals onto meanings and vice versa. In this framework, the only thing that can go wrong is - as in Shannon's theory - noise that disrupts the transmission of symbols (e.g. phonemes) which then obviously also disrupts the subsequent interpretation of the signal, because if the phonemes are wrongly recognized, the words and sentences will be wrongly recognized as well. While this is indeed one of the challenges any communicator has to meet, it is not the only challenge, and probably not the biggest one either. Indeed, having a bad phone connection when we talk to other humans, or having too much ambient noise for automatic speech recognition to operate reliably is disruptive to communication.

A problem that is much harder to cope with is the *pragmatics problem*: in human communication, there is no 1:1 mapping between signals and their intended meanings. This is because the meaning of a communicative signal is not entirely fixed and context free, but depends on the *communicative intention* of the generator of that signal in a particular context (i.e., given a particular speaker, addressee, physical setting, and discourse history). The linguistic discipline of pragmatics is concerned with exploring this problem and its possible solutions, and the pragmatics literature is littered with examples showing that the mapping of signals to communicative intention is not 1 to 1, but rather Many to Many [3-6]. That is to say, every signal can convey many different communicative intentions, and every communicative intention can be conveyed by generating many different signals. For example, the phrase "thank you" can be intended as the expression of gratitude (when accepting a gift), the expression of annoyance (e.g. ironically, after someone has just dropped his coffee in your lap), or even as a request (a tennis umpire trying to quieten down the spectators during a match). However, the tennis umpire can also try to quieten down the spectators by saying "please be quiet", "quiet please", or "shhhhhht!". To complicate things further, communicators not only use language to communicate, but also facial expressions, gestures, eye-gaze, and functionally motivated actions (e.g. placing the book one wants to buy on the counter of the shop). For these nonverbal signals, it is even harder to establish an unambiguous meaning associated with it, as there are no "dictionaries" to guide us.

In this paper, we outline a blueprint (which we call 'the BIC model') for a real communication module, one that incorporates the pragmatics problem by making it central. The model is intended to serve as an architectural guideline for communication modules in artificial systems, but is also sufficiently general to use as a cognitive-psychological model of communication between any possible combination of humans, animals, and machines. The core assumption of the model is that for real communication, there are three distinct and essential components needed: a set of "biological" or "built in" constraints (the B-system), an interaction engine or social-interactional cognition component along the lines of [7] (the I-system), and a

content-addressable database containing conventions (the C-system). We now discuss the components of the BIC model, its implementation and some of its applications.

2 Biological or Built-in Constraints

Whether biological or artificial, an agent that communicates has physical constraints that limit the type of signals that it can produce and perceive. Robots have cameras with a certain resolution and sampling rate, humans have eyes with specific perceptual limitations (e.g. the visible colors). Robot use microphones, while humans have ears. Both can detect only a limited range of frequencies. These constraints limit the possible signals, with consequences for the way communication proceeds. In humans, the B-system also represents whatever communicative intelligence is innate, and is responsible for the human infant to start 'bootstrapping' into communication (e.g. by responding to certain frequency patterns, recognizing its mother's face, etc.). This is necessary, as the C(onvention)-system, discussed below, is assumed to still be empty in the newborn infant. In a communication device a la Shannon (e.g. a computer modem) the B-system is the capacity to distinguish 1s and 0s, leading to the capability to receive a sequence of 'bits'. The same holds for production: a robot has a number of output devices (a speaker, a head, a face, hands, etc) just as humans have faces, a voice, eyes etc. Even in the simple case of the Shannon device, a modem should be capable of creating 0s and 1s to send messages to the other side.

The anatomy of the human body, including its articulatory organs, is subject to constraints that are determined by the genetic make-up of homo sapiens. In the case of Robots, these constraints are in its physical design, and ideally these are as similar as possible to the human constraints, but this is generally not the case. While a range of artificial sensors exist that are more sensitive and accurate than the human receptors, there are as yet no artificial effectors more dextrous and flexible than the human hand.

3 The Interaction Engine

The Interaction Engine, or I-system, is a processor that is central in the BIC model. It has three main tasks. First, it performs the computations necessary to attribute a communicative intention to the sender of received signals (delivered to it via the B-system). Second, it performs the computations necessary to generate signals that will enable others' I-systems to successfully attribute communicative intentions. These two functions are essential, because of the pragmatics problem mentioned in the introduction: the Many to Many relationship of signals to communicative intentions. A third and important additional function of the I-system is to manage the C-system (the database of stored conventions), which will be discussed below. Importantly, the I-system is modality-independent; it will process a verbal utterance, the raising of an eyebrow, or gestures that accompany speech (see [8]) in the same way.

There have been different proposals about how to design an I-system. In the dialogue processing model of Pickering & Garrod [2] the I-system is assumed to be unnecessary most of the time. It only kicks into action when special conditions apply,

such as persistent misunderstandings or deliberate deception. According to Pickering and Garrod, the reflexive computations that are presumably involved in solving the pragmatics problem are computationally too expensive to be constantly governing ordinary communication.

In the approach to the pragmatics problem called *relevance theory* by Sperber & Wilson [6] it is assumed that the correct form/meaning mapping can be extracted by giving precedence to inferences that maximize 'relevance' while minimizing the cognitive effort needed to compute it. In contrast to Pickering and Garrod, Sperber and Wilson do claim that some type of inferential processing is necessary for communication, but they differ from the Gricean approach [3, 9] in the way this processing is performed. They assume that Grice's maxim of Relevance is central, and sufficient.

The approach we take for building an I-system is essentially neo-Gricean, following Levinson [5, 7]. We propose that the I-system solves the pragmatics problem by using conceptual *simulation* of the interlocutor(s). In short, senders of signals simulate how a receiver would interpret the signal, and receivers simulate how a sender could have decided to send that particular signal. Recent neurocognitive findings from the EU funded project JAST have shown that communicators solving the pragmatics problem while sending signals activate the same well-identified brain region as receivers do while solving the pragmatics problem for comprehension [10]. We call this type of simulation *conceptual* to distinguish it from *sensori-motor* simulation, as has been proposed by researchers who study "mirror neurons" [11]¹. In a conceptual simulation, what is simulated are the *inferences* of the interlocutor, given a certain set of propositions about the world (i.e., his beliefs and desires, see e.g. [12]) and a certain signal that is received or being sent. Such a simulation can never be perfect, as information about the beliefs and desires of others is always incomplete. However, propositions of which the truth value is both unknown and essential for the inferences at hand are assumed to be copied from the communicator's own beliefs and desires. (In any case, it is not critical to have an exact copy of your interlocutor's mental state, but to have one that is 'sufficient for current purposes'; Clark [13]) This is a computational *heuristic* that leads to a degree of 'egocentrism' in communication. Evidence suggests that communicators often do not take current common ground among interlocutors into account, for instance in producing and understanding referring expressions [14-17]. While communication proceeds, participants gradually build up a more and more accurate representation of the other (or, to put it in other words, they will develop more and more *common ground*, see [18]) which enables them to take their interlocutor's perspective into account. This process is beautifully illustrated in an experiment by Clark & Wilkes-Gibbs [19], in which communicators are shown to gradually establish so called *conceptual pacts*, which are essentially form/meaning mappings that have been implicitly agreed upon by communicators. Updating our model of interlocutors during the course of communication has been called *grounding* by Clark & Schaefer [20].

While we propose to build an I-system based on mutual simulation, we appreciate the point by [14-17] and others that reflexive computations can be expensive. To

¹ In the fMRI study reported in De Ruiter et al., there was no activation in the mirror neuron areas while communicators were trying to solve the pragmatics problem.

avoid this resource bottleneck, the I-system can store the results of its pragmatic simulations in the C-system, which constitutes a content-addressable database of recurrent form/meaning mappings. Take for example the following short dialogue:

- A: What did you think of Steve's latest book?
 B: It put me to sleep.

The first time A encounters B's response, she might well need to engage in conceptual simulation to extract the communicative intention from the speech "It put me to sleep" (as the 'literal' meaning does not make much sense), but after having inferred (and validated, through grounding) the communicative intention of the utterance, the mapping of the signal "X put me to sleep" onto the communicative intention "I want you to know that I found X rather boring" can be stored for later use. Thus, we now discuss the third and final component of the BIC model, the C-system.

4 Conventions

Efficient communication between agents is impossible without the ability to rely on socially shared *conventions* about what signals mean. For example, most if not all implemented models for linguistic communication contain a lexicon, a database that stores the relationship between the surface form of words and their associated meanings. However, storing conventional form/meaning mappings need not be limited to words. It can also be used to store mappings between phrases and their associated communicative intention. For example, the phrase "How are you?" is frequently not a request for information, but a greeting, and storing this mapping can save the I-system a lot of processing cycles. Similarly, parameterized *templates* could be stored as well. The form "Could you <VP>?" is most often a polite request to VP. Again, storing this mapping avoids having the I-system compute the communicative intention from scratch every time it is encountered.

Syntactic rules also establish a system of shared rules that govern the ordering of and mutual relations between linguistic symbols. At a more peripheral level of processing, the same holds for phonology. In fact, traditional linguistics has mainly focused on charting out the conventions governing linguistic competence. However, we see the C-system as a place to store not only the general linguistic *competence* of the language user, but also aspects related to performance, or language use. For instance, the *conceptual pacts* studied by Clark & Wilkes-Gibbs [19] can be treated as local conventions, conventions that only apply for a certain context and with a certain interlocutor. A good example of local conventions can be found in the domain of person reference [21]. In a limited social group where everybody is assumed to know him, an individual can be referred to as "Herb", but in the wider context of a psycholinguistics conference, the reference "Professor Clark" might be needed to achieve recognition. The localized nature of conventions represents a challenge for the implementation of the C-system. Conventions have a "social scope" which needs not only to be represented in the C-system, but in fact provides an essential *retrieval key*. If I want to refer to a person, I do not want to look through all the social groups I am a part of in sequence before deciding on a reference: I want to be able to index the retrieval of the reference by the social group or individual I am designing my person

reference for. A similar argument can be made for the *context* of communication. When I'm socializing in a bar with my friends, other conventions apply than when I am in a formal meeting at work, even if I am in that meeting with the same people. By using both the social scope and the social context as a retrieval cue for the C-system, the I-system can retrieve those conventions that are 'relevant' to the situation at hand first, before considering other conventions.

To summarize, when the I-system receives a signal via the receptors of the B-system, it first checks the C-system to see if that signal with that context (or parts thereof) has been stored. If it has, the I-system will use that stored mapping as a basis for attributing a communicative intention to the signal. If it is not, the I-system will engage in simulation of the interlocutor, and try to compute the communicative intention. If this turned out to be the correct interpretation, it will store the newly found mapping into the C-system. Analogously, if the I-system needs to produce a signal to get across a certain communicative intention, it will first look in the C-system to see if it contains a previously stored meaning-form mapping that can be used. If not, it will engage in simulation (of the receiver) to generate a signal that might work. If it turns out to have worked, this new mapping will also be stored in the C-system for later work. To optimize retrieval efficiency, the C-system is a content-addressable database that allows social scope and context to be used as retrieval keys.

5 Summary

In this paper we have outlined a blueprint for a communication module (see Figure 2). It is based on three essential components: (1) a set of biological (hard-wired) constraints that define the properties of the possible incoming and outgoing signals, called the B-system, (2) an interaction engine that performs the inferential processing

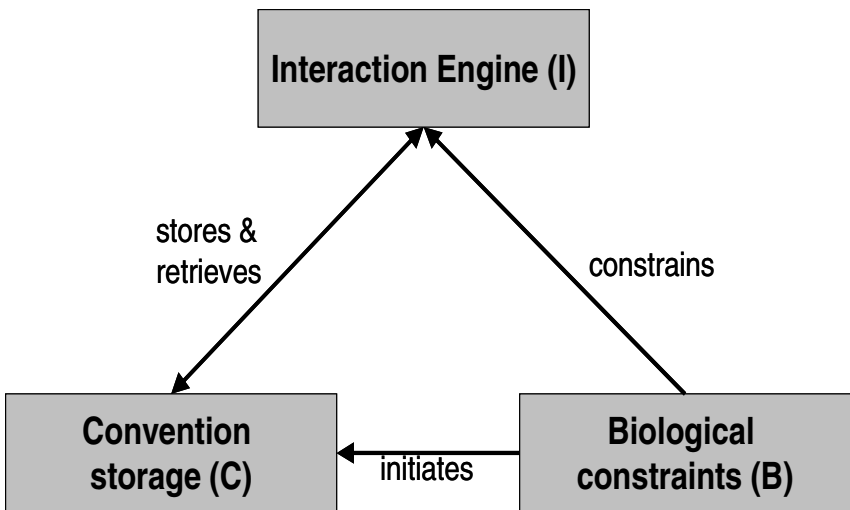


Fig. 2. General layout of the BIC architecture for agent-agent communication

needed to solve the pragmatics problem of mapping communicative intentions to signals and vice versa, and (3) a database of socially shared communicative conventions about the mappings between signals and communicative intentions. The more "experience" the system will have, the more sophisticated and efficient it will operate. For the implementation of the I-system, we propose to use neo-Gricean mutual simulation, combined with rational use of the C-system to take shortcuts whenever possible.

The BIC model can be used both as a conceptual model in the study of agent-agent communication (including HCI), and as a blueprint for an implementation of a communication module in artificial agents. This way, insights about communication between agents can flow in both directions. Challenges arising from implementing a communication module in agents can stimulate interesting new research issues in human-human communication, and existing knowledge of the latter can be used to improve the quality of our implementations in artificial agents.

References

1. Shannon, C.E.: A mathematical theory of communication. *Bell. System Technical Journal* 27, 379–423 (1948)
2. Pickering, M.J., Garrod, S.: Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(2), 169–226 (2004)
3. Grice, H.P.: Logic and conversation. In: Cole, P., Morgan, J.L. (eds.) *Syntax and semantic 3: Speech acts*, pp. 41–58. Academic press, New York (1975)
4. Levinson, S.C.: *Pragmatics*. Cambridge University Press, Cambridge (1983)
5. Levinson, S.C.: Interaction biases in human thinking. In: Goody, E.N. (ed.) *Social intelligence and interaction*, pp. 221–260. Cambridge University Press, Cambridge (1995)
6. Sperber, D., Wilson, D.: *Relevance: communication and cognition*, 2nd edn. Blackwell, Oxford, UK (1995)
7. Levinson, S.C., On the human interactional engine. In: Enfield, N.J., Levinson, S.C. (eds.) *Roots of Human Sociality: Culture Cognition, and Interaction*, Berg: London (2006)
8. De Ruiter, J.P.: Postcards from the mind: the relationship between thought, imagistic gesture, and speech. *Gesture* 7(1), 21–38 (2007)
9. Levinson, S.C.: *Presumptive Meanings; The Theory of Generalized Conversational Implicature*. The MIT Press, Cambridge, Massachusetts (2000)
10. De Ruiter, J.P., et al.: On the origin of intentions, in *Attention & Performance XXII*. (in press)
11. Gallese, V., Goldman, A.: Mirror neurons and the simulation theory of mind reading. *Trends in Cognitive Science* 12, 493–501 (1998)
12. Jurafsky, D.: Pragmatics and computational Linguistics. In: Horn, L.R., Ward, G. (eds.) *The Handbook of Pragmatics*, pp. 578–604. Blackwell, Oxford (2006)
13. Clark, H.H.: *Using language*. Cambridge University Press, Cambridge (1996)
14. Bard, E.G., et al.: Sharing the cognitive burdens of dialogue. (in press)
15. Barr, D.J., Keysar, B.: The paradox of egocentrism in language use. In: Colston, H.L., Katz, A.N. (eds.) *Figurative language comprehension: Social and cultural influences*, pp. 21–42. Erlbaum, Mahwah, NJ (2005)
16. Keysar, B., Barr, D.J., Balin, J.A.: Definite Reference and Mutual Knowledge: Process Models of Common Ground in Comprehension. *Journal of Memory and Language* 39, 1–20 (1998)

17. Horton, W.S., Keysar, B.: When do speakers take into account common ground? *Cognition* 59, 91–117 (1996)
18. Clark, H.H., Marshall, C.R.: Definite reference and mutual knowledge. In: Joshe, A.K., Webber, B.L., Sag, I.A. (eds.) *Elements of discourse understanding*, pp. 10–63. Cambridge University Press, Cambridge, U.K (1981)
19. Clark, H.H., Wilkes-Gibbs, D.: Referring as a collaborative process. *Cognition* 22, 1–39 (1986)
20. Clark, H.H., Schaefer, E.F.: Contributing to discourse. *Cognitive Science* 13, 259–294 (1989)
21. Enfield, N.J., Stivers, T. (eds.): *Person reference in interaction: linguistic, cultural and social perspectives*. Cambridge University Press, Cambridge (2007)