Cognitive Interaction Analysis of Clinical Encounters

Wayne Zachary, Ph.D and Russell Maulitz, M.D., Ph.D
CMZ Health Technologies LLC
Fort Washington, Pennsylvania, USA

Cognitive task analysis (CTA) is a standard HFE method that assumes a tool-use context involving clear situational goal(s). Healthcare contexts involve patient-provider interactions that can render CTA inappropriate. Patient and provider often have differing or conflicting goals, and have their interactions complicated by communicative barriers arising from cultural and even personality differences. Cognitive interaction (COIN) analysis is introduced as a new formal method for decomposing and documenting provider/patient interactions for human factors purposes (using cancer-care examples). It represents the interaction as a story graph, and each party in the interaction as a social agent with individual motives, personality/cultural filters, and knowledge.

INTRODUCTION

Cognitive Task Analysis (CTA) refers to a family of methods (Schaapen, Chipman & Shalin, 2000; Crandall, Klein, & Hoffman, 2006) that are widely used to cognitively decompose a work-related activity (i.e., task) undertaken by one or more people employing computer-based or mechanical tools to achieve specific work-related goals. CTA has become a core Human Factors Engineering (HFE) method for decomposing and documenting tasks involving human information processing, as a foundation for modeling and engineering analysis (e.g., Zachary, Ryder, & Hicinbothom, 2000), design (e.g. Ryder, Szczepkowski, Weiland, & Zachary, 1998), and evaluation (e.g., Kiers & Santoro, 2004; Saitwal, Feng, Walji, Patel & Zhang, 2010) of human-machine systems.

As HFE has become more involved in health care, CTA has come along with it (e.g., Kushniruk, 2001; Weir, Nebeker, Hicken, Capo, Drews, & LeBar, 2007). In those aspects of health care where the engineering problem deals with human interaction with devices, this can be an appropriate application. However, the practice of medicine is certainly at the heart of healthcare, and much of the practice of medicine involves clinical encounters -- direct interactions between doctors and patients. The clinical encounter is not a cognitive task, even in the broadest application of the above, highly general, definition.

The encounter involves two people interacting (not a person and a machine), and does not involve a common or shared goal that defines task accomplishment. Indeed, an encounter can be completed with the two parties having distinctly different viewpoints on whether it was successful or not. The patient, for example, often has a goal of gathering as much information from the physician as needed to address his or her concerns, fears, questions, etc., regardless of time. A the same time the physician may have a competing goal of keeping the interaction within a general time-frame associated with the type of encounter.

The patient and the physician thus each have different interests in the interaction. Neither person is in control of the interaction, in the sense in which the driver of a car, pilot of aircraft, operator of a machine (even a complex machine with embedded intelligence) is in control of the interaction. This is important because CTA, in all its various forms, begins with focusing on the goal(s) or decision(s) that guide the person’s reasoning and actions in completing the task, and then decomposes the knowledge and reasoning used to achieve the goal(s). Thus, because the clinical encounter is fundamentally a dyadic interaction, not a cognitive work task, CTA is inappropriate or at best insufficient. Here, we present a new technique for the cognitive front-end analysis of clinical encounters, called Cognitive Interaction (COIN) analysis. COIN Analysis is rooted in conventional CTA but integrates social psychology constructs for analyzing and modeling human interactions.

We note that CTA (or a specific CTA method) can be defined in two distinct and complementary ways. First, it can be defined in terms of a representational product that exposes the internal information manipulation processes typically utilized by a person with a given degree of experience (often an ‘expert’) with that task being performed. Second, it can be defined in terms of the data collection and manipulation process that is used to create the (CTA) representation for a specific task (and level of experience). We define COIN analysis here only in the first way, in terms of the representational format that it uses. However, some discussion is provided in the conclusions on the process that can be used to construct a COIN representation.

COIN ANALYSIS FORMAT AND REPRESENTATION

Like CTA, COIN assumes that both parties engaged in the interactional task are sufficiently familiar with the situation that they have developed a mental representation of that interaction, and use that mental representation to generate their side of the interaction. COIN analysis derives from a family of socio-cognitive methods used in social psychology and social anthropology to capture the cognitive structure of a dyadic interaction (Schank & Abelson, 1977; de Beaugrande & Colby, 1979; Lehnert, 1981; Miller & Read, 1991; Graesser, Singer and Trabasso, 1994; Zachary, Read, Miller, & Santarelli, 2010).

The representation contains two hierarchical levels. In the top level, the general structure of the interaction is identified via distinct roles interacting within a specific situational context via sequential verbal/communicative acts. These roles and sequence of acts are depicted as a graphical network called a basic Narrative Structure Graph (NSG). In the second
and more detailed level, the basic NSG is annotated and elaborated to create an enhanced NSG incorporating the effects on the interaction afforded or promoted by specific individual characteristics of interest. These can include personality traits, cultural values, and situational motives that affect the situated cognition and behavior of different individuals that may assume the roles in the interaction (i.e., the basic NSG). Thus the basic NSG identifies a network of ways in which a specific interactional situation might play out, and the enhanced NSG identifies how different individuals might make specific interactional choices arise within the general space of interaction.

To make the abstract definition of COIN analysis more concrete, we use an example of an interaction between a physician and a patient in which the physician is attempting to discuss the options and implications of prostate cancer screening tests (obviously making the patient a male, and justifying the use of masculine pronouns to refer to the patient role where appropriate).

**Defining Interaction Structure**

The top level of a COIN analysis decomposes and formally represents the range of interactions in a clinical encounter (or more generally, any purposeful interaction). The interactional dynamics are captured in the NSG, which depicts the tacit understanding of the ways in which the interaction may proceed or evolve as that understanding is shared between the participants in the interaction. The importance of the understanding defines a level of experience or familiarity required for COINS analysis as analogous to that of linguistic competence (Chomsky, 1965). That is, each participant is presumed to have interacted in a given role in this type of interaction, in the local cultural context, sufficiently frequently to have internalized the underlying structure of the interaction to the point that they can use the understanding to guide their behavior.

An NSG decomposes the encounter into sequential transactions in which patient and physician (or more generally, the participants) take turns speaking and listening. The transactions are abstracted into generic types called narrative units. Each narrative unit contains an indicator of *role* in the interaction, a label summarizing the intention of the *communicative event* from that role’s viewpoint, and the *perceptual environment* that the communicative event offers to the other role(s) in the interaction. The additional concepts (in italics) introduced above are defined in the following paragraphs.

A *role* is a social construct with a shared social/cultural understanding that a person assumes in an interactive situation. In the clinical encounter, there is a physician role and patient role. A person may act in the patient role even if, outside the encounter, that person is a physician, or has some other role with regard to the individual in that physician role. The point is that, within this interaction, the two interact via their roles, and according to their shared understanding of how those roles interact (i.e., according to the NSG).

*Communicative events* are pragmatic constructs that describe verbal (and paralinguistic) actions a person in one role uses to communicate a single message to the person in the other role. It could be as simple as a single word or phrase (e.g., “difficulty going to the bathroom”) or as complex as a larger multimodal set of utterances, gestures, facial expressions, tone of voice or body language.

The *perceptual environment* identifies the key cues that are perceptually available and salient to the patient and physician in each narrative unit. In the clinical encounter, communicative transactions are primarily verbal, so the utterances of one agent are the primary cues available to the other, although the tone and timber of the agent’s voice are also cues that are available. During the verbal exchanges that make up the bulk of the encounter, each participant also has available the paralinguistic expressions of the other, including facial and body language. These may be processed to infer cognitive and emotional states (such as engaged/distracted, concerned/unconcerned, worried/not-worried). For example, during a discussion in which the patient is trying to press questions about some element of the encounter (e.g., diagnosis, prognosis, treatment plan), a movement of the physician toward the exit from the examining room or placing a hand on the door knob of the examining room is a cue that may be perceived as an intention to cut short the discussion and end the encounter, and further, an inference that the physician is distracted and unconcerned.

A Narrative unit (NU) is drawn as shown in Figure 1, as a frame with slots for the role, the communicative event and intent, and the perceptual environment to the other role(s). Each NU is also given a unique number in the NSG.

![Figure 1. Narrative unit Representation](image)

Because the NSG presumes a turn-taking interaction, NUs containing communicative events by one role (e.g. Physician) transition to NUs containing communicative events by the other. For example a simple interaction to initiate a dialog once the Physician enters the examining room might connect a communicative event to greet the patient as how she/he is with at least two canonical responses, one in which the patient responds with a chief complaint, and another in which the patient responds with a more generalized response (i.e. “small talk”) to engage the physician in brief interaction before responding at some point with a chief complaint, or before the physician again tries to steer the interaction toward the clinical reason for the visit. This is shown in the example in Figure 2.

A cycle of NUs can be created via a transition arrow that connects back to a prior NU of the other role. In Figure 2, a cycle of small talk is created that way, and can continue indefinitely, until one of the roles moves it toward a different subject. Figure 2 also shows that repeated components of the perceptual environment can be externalized and referred to in the NU diagram to simplify the appearance.
When an NU contains more than one existing transition arrow, it indicates a situation where the other role may respond in multiple ways to that NU, creating an explicit or implicit decision point for the other role. When the NUs are augmented in the second level COIN analysis, additional detail is added to indicate how individual differences of various kinds influence such decisions.

Figure 2. Example NSG for Initial Stage of Encounter

There can be sub-networks of an NSG that deal with a segment of the interaction that the participant would understand as related, such as the unit shown in Figure 2 above. Such sub-networks can be grouped as a Story Unit, allowing that sub-network to be treated as a single block or piece of the overall space of the interaction. Figure 3 shows this, with the Story Unit grouping NUs 1, 3, and 4. To complement Figure 3, a separate diagram showing NUs 1, 3, and 4 and their transitions would be needed.

Figure 3. NSG Showing Story Units and Narrative Units

Figure 4a shows an example NSG for a full type of encounter, based on the example case of an annual physical examination for a middle-aged man in which the physician plans to bring up the subject of screening for prostate cancer. After the initial interaction to engage the patient (as in Figure 3), the interaction follows a nominal path beginning by a one-by-one review of systems.

This nominal path is depicted along the center path in the figure (which is rotated to fit in a single page clearly). The cycling arrows show that the same general interaction is used for each physical system, and that each may produce a pertinent negative or positive response from that patient that may take the interaction into a more detailed and specialized question/answer and possibly a more focused physical examination.

After this general examination is complete the physician decides to start the “cancer talk” (abbreviated as CA-talk). In this CA-talk, the physician reviews the relevance of the disease to the patient’s age and specific risk factors (e.g., family history, belong to high-risk populations, etc.) and describes the suggested strategy of including a Digital Rectal Exam (DRE) accompanied by blood testing for PSA (prostate specific antigen) levels. As this proceeds, the interaction may get more complex. The patient can simply accept the recommendation and proceed, or he can interrupt with questions (or save them for his ‘turn’) of various kinds. These could be expressed as medical questions, as reluctance to undergo the DRE, or seemingly unrelated social or cultural issues.

This is the point where there are opportunities for miscommunication compounding the interaction process. An individual in either role can fail to understand the previous communicative event because that person is not attending to all modalities of the communicative act. This is especially true for the physician role, as the doctor may not recognize the social or cultural issues being voiced by the patient, and thus only responds as if they were medical concerns, leading to emotional discomfort and often disengagement on the part of the patient (see Santarelli, Maulitz, Zachary, Barnieu, and O’Connor, 2009). Through the multiple interaction paths involved here, the patient may accept the initial recommendation, or refuse it, in which case the physician assumes a fall-back recommendation for a colonoscopy (assuming there is no recent one) and PSA testing without the DRE. Again through discussion this choice may be accepted or the colonoscopy alone may be accepted, or all options may be deferred (probably to be raised again in next year’s annual physical).

Figure 4b shows a disconnected sub-NSG representing an interaction segment in which the patient has a pressing question or problem that is deliberately undiscussed until some point in the interaction in which the patient brings it out. This ‘surprise’ sub-interaction is not uncommon, but leads to an otherwise out-of-place discussion of this problem or issue in terms of the general flow of the interaction narrative.
(PE1) = Physician’s utterance (words, syntax, tone, accent), facial expression, body language, positioning in encounter room, anything physician is holding; appearance & content of encounter room

(PE2) = Patient’s utterance (words, syntax, tone, accent) Facial expression, body language, clothing, other persons brought into examining room

Figure 4a. Example Basic NSG for Annual Check Involving Prostate Cancer Screening Discussion

Figure 4b. Sub-narrative graph for Denial Narrative
Cognitive Task Analyses, such as the many variants based on the GOMS notation (Card, Moran, Newell, 1983), and even more traditional task analysis representation like task network models (Laughery, Archer, Plott & Dahn, 2000) provide a generative representation of the space of actions that a human operator can take in a person-machine environment under different problem conditions. This generative property is maintained in the COIN analysis via the NSG. The specific sequence of choices made by both participants creates a specific path taken through the interactive space defined by the NSG. Each such path constitutes a post-hoc narrative account that tells the story of that specific interaction.

In additional to specific narrative accounts, the concept of narrative can be used to capture typical or canonical encounter evolutions. For example:

- The linear or nominal narrative defines the ideal or expected encounter. The path of steps down the center of the NSG defines this nominal narrative. If this is the only narrative of the encounter that is used to evaluate clinical processes, or to design health information technology (such as encounter documentation tools) intended to support the encounter process, then most of the actual variability in the encounter will be missed.

- The surprise narrative, in which the patient has a hidden agenda that emerges at an unforeseen moment in the interaction, is another evolution frequently seen by physicians. This involves the disconnected ‘surprise’ sub-narrative shown in Figure 4b.

- Another common narrative is that of the resister or questioner, a patient who questions actions, queries, suggestions, by the physician, and often resists some or all advice. Increasingly, the questioning and resisting is supported (on the patient’s part) by information read or downloaded from the broad internet.

- The social/cultural issue narrative occurs frequently when the patient and physician have dissimilar cultural backgrounds, or when the physician does not understand the patient’s social situation. This narrative can often appear as a ‘resistance’ narrative if the physician does not recognize the cultural issue involved, and deal with it directly and appropriately (see Santarelli, et al, 2009).

One final common narrative that is not included in Figure 4 is what can be called a denial narrative. In this evolution, the patient hides or denies symptoms or behaviors (e.g., drug use, alcohol use, sexual behaviors) that the physician needs in order to understand and develop an testing or treatment plan for other observed systems or problems. To capture this narrative, additional narrative units would be needed to reflect the interactions between the doctor and patient to either uncover the denied elements or to find some other way to get at the underlying problem without the patient’s admission of the information being hidden.

**Augmenting Interaction Structure with Individual Differences**

While the basic NSG that is produced by a COINS analysis identifies the multiple narrative paths that can arise from a specific type of encounter, it does not provide insight on why these different narratives occur in different instances of the interaction, nor does it provide information on the features of the agents or of the environment that lead to one individual narrative state transition to occur instead of another. These aspects are addressed in the augmented NSG, the final step in COIN analysis.

It should be noted that in many cases the basic NSG will be sufficient for the human factors activities that may ensue. For example, just knowing the different narrative paths in an encounter may be sufficient to allow an appropriate level of flexibility in a control mechanism for an information system intended to support recording the encounter. In other cases, however, the understanding of the circumstances in which one specific narrative evolution may occur (instead of another) can be vital to succeeding human factors activities. Designing training scenarios or evaluating physician clinical skills are examples of such activities.

In this discussion of the augmented NSG, we focus in particular on how individual differences are captured in COINS analysis. Before introducing the augmented NSG, it is useful to reconsider the meaning of transitions in a basic NSG. Figure 5 shows a sub-graph in which an NU representing a physician’s communicative act can be followed by one of two different NU types. In the following discussion, we denote the two transitions in Figure 5 as (J,K) and (J,L).

It is important to understand that the structure in Figure 5 means that it is the patient who chooses the succeeding NU options (either K or L), not the physician. COINS analysis allows the analyst to view the choice among potential successor NUs in terms of the affordances that the different NUs provide for the agent to pursue situational goals and/or to express other individual characteristics.

![Figure 5](Image)

**Figure 5.** Patient chooses NU K or L, not the doctor

In an augmented NSG, specific transitions can be annotated to show the differential relevance of that transition to various characteristics of the agent acting in the specified role in a specific instance of the interaction. The characteristics reflect features of individual differences, and can be theory-based or analysis-case specific. Examples of general, theory-based features include:

- **Motivations** are more pervasive than situational goals, motivations represent fundamental social orientations, (e.g., establishing dominance, avoiding social separation or isolation, social or filial bonding, etc.). Different NU options afford different opportunities to pursue specific motivations. Research in social psychology has argued strongly that combinations of such general motivations form the substructure for individual personality traits (see
Miller & Read, 1991; Read, Monroe, Brownsetin, Yan, Chopra & Miller, 2010).

- **Emotions** may clearly change during the encounter, and appraisal theories of emotion (e.g. Roseman, 2001) argue that emotions reflect the perceived relationship between underlying motives and goals and projected possible future states. Different NU options may afford greater or lesser attractiveness to avoidance of negative emotions or pursuit of more positive ones.

- **Beliefs or cognitive filters**, often developed from participation on specific cultures and/or subcultures, are reasoning biases that lead individuals to frame actions or others in a specific way. This framing can, in turn, give some NU options greater affordances for pursuing that filtered view. For example, patients without insurance may believe that their care is inferior as a result, and incorporate that belief into their interaction strategy with the physician (see Santarelli, Maulitz, Zachary, Barnieu, & O’Connor, 2009 for an example of the use of belief filters in the context of a COIN-like analysis).

The simplest way to create an augmented NSG is to determine the characteristics of interest and a set of categories for each. For example, consider the transitions (8,9) and (8,12) from Figure 4a above. This is the ‘cancer-talk’ part of the encounter, in which the physician tries to encourage the patient to agree to a DRE and a follow-up colonoscopy. Figure 6 shows example simple annotations to the basic NSG format that turn it into an augmented NSG representation.

![Figure 6. Example Transition Annotations](image)

In Figure 6, both transitions have been annotated with affordances for emotional state and underlying motives. The (8,9) transition, corresponding to the patient’s agreement to the physician request, shows that it would increase a patient’s fear (of a negative finding) and a potential strong disgust with the mechanisms of the DRE itself. It also shows an affordance for the broad motivation of self-preservation, as that patient’s concern for avoiding prostate or rectal cancer or encountering an early/treatable form would appeal to the desire for long life.

In contrast, the (8,12) transition has been annotated with an emotional valence for relief, indicating an affordance to feel relief by avoiding the uncomfortable exam and feared possible negative result. It has also been annotated with an underlying motivation to avoid threats or harm, based on a patient’s fear of pain in the procedure and the threatening potential of cancer in general.

The example annotations, in the case of Figure 6, are qualitative, showing only the opportunity for emotions or motives to be increased or decreased in activation. Such annotations need not be applied to all transitions in an NSG, nor do then need to be qualitative. The analyst can attempt to quantify them, although in such cases it is recommended that the quantification be done on a normalized scale of [0,1].

The use of the annotations lies in understanding how the affordances they capture will make different NU options more or less attractive to individual actors with different personality traits, differential emphasis on underlying motives, different underlying beliefs and cognitive/cultural filters, etc. Most of these characteristics will themselves vary in degree of activation throughout an interaction, creating interactions between baseline or beginning activations of individual characteristics, and activation states at various points in the encounter.

To deal practically with such fine-grained and complex interactions, it may be necessary to use the augmented NSG to construct computational models. Such models can represent the individual characteristics of specific agents that are interacting and can simulate the interactions dynamically, calculating the changing internal state of each agent to best predict or model which specific choices will be taken through the interaction. A key aspect of such agent-specific computational models is the ability to model the changes in activation of relevant individual factors as the interaction proceeds. Recent research (Marsella, Pynadath, & Read, 2004; Read, Miller, Kostygina, Chopra, Christensen, Corsbie-Mason, Zachary, LeMentec, Iordanov, & Rosoff, 2008; Deaton, Barba, Santareli, Rosensweig, Souders, McCollum, Seip, Knerr & Singer 2005) produced a number of models that are able to do this. Such computational systems could be used to translate analyses from this level of COINS analysis into executable simulations.

**CONCLUSIONS**

In this paper we have introduced cognitive interaction or COIN analysis, a new methodological tool for front-end analysis and data collection in human factors. It has both broad similarities and specific differences and advantages over cognitive tasks analysis, which has become well-established in the field. Like CTA, COIN analysis:

- explicitly decomposes a process into discrete but abstract units called Narrative Units;
- utilizes hierarchical constructs called Story Units that allow complexity to be managed;
- encapsulates cognitive components of intention, perception and external cues and uses these to create an account of how a person frames and selects actions in a specific problem context;


- Is generative, providing a single abstracted representation that allows multiple instances of behavior to be captured as different paths through the representational space.
- Can be used, because of its generative property, to create:
  - use-cases for engineering design of artifacts such as diagnostic aids, EHR encounter-capture interfaces or physical facilities;
  - ranges of case narratives for evaluation of a design, of a provider’s skill in using a new artifact, or a provider’s skill in clinical communications;
  - complementary scenarios for scenario-based training and/or for simulation/game-based training systems;
  - development of computational cognitive models.

However, COINS has multiple features that CTA lacks when the focus of analysis is an interaction between patient and provider. Specifically, COIN analysis:

- brings agency to both provider and patient roles in the encounter, treating both parties as having intentions in the interaction that can differ and at time conflicts. The focus of CTA on a specific individual does not readily allow for these differences in interests and intentions;
- is broadly applicable to other interpersonal work domains within healthcare (and beyond). While this paper has focused on the clinical encounter, COIN analysis is equally applicable to other clinical interactions that can and do occur at key points in the health care process, and even beyond health care to other domains of interest in human factors;
- provides extensions to account for individual differences at the level of personality, motivation, and culture, while CTA typically considers at most differences in perceptual/motor skills or abilities.

COIN analysis, because of its use of the Narrative Structure Graph, implicitly makes the large body of literature on narrative and discourse analysis relevant and applicable to the human factors analysis of interactions. Among the limitations of COIN analysis is the lack, at present, of (software-based) tools to support specifically the development and manipulation of COIN analyses (although development of such tools the subject of future research of the authors).

REFERENCES


