Establishing Causality in Complex Human Interactions: Identifying Breakdowns of Intentionality

Peter Goodison[1,*], Peter Johnson[1], and Joanne Thoms[2]

[1] Department of Computer Science, University of Bath,
Claverton Down, Bath, UK
{p.goodison,p.johnson}@bath.ac.uk
[2] BAE Systems Integrated System Technologies Limited,
Lyon Way, Frimley, Surrey, UK
Jo.thoms@baesystems.com

Abstract. People in complex scenarios face the challenge of understanding the purpose and effect of other human and computational behaviour on their own goals through intent recognition. They are left asking what caused person or system 'x' to do that? The necessity to provide this support human-computer interaction has increased alongside the deployment of autonomous systems that are to some degree unsupervised. This paper aims to examine intent recognition as a form of decision making about causality in complex systems. By finding the needs and limitations of this decision mechanism it is hoped this can be applied to the design of systems to support the awareness of information cues and reduce the number of intent recognition breakdowns between people and autonomous systems. The paper outlines theoretical foundations for this approach using simulation theory and process models of intention. The notion of breakdowns is then applied to intent recognition breakdowns in a diary study to gain insight into the phenomena.

Keywords: Intentions, Decision-making, Awareness, Autonomous Systems.

1 Introduction

People have a fundamental need to understand the world around them. Part of that understanding is reasoning about cause and effect to aid the prediction of future events in the environment. Due to the competitiveness of organisms, being able to make decisions faster, especially about the future actions of other organisms is advantageous. Such was the benefit of predicting future states to our ancestors, information gathering, decision making, problem solving and anticipatory judgment processes and mechanisms evolved to improve the speed and accuracy of this type of foresight. With this ability humans can recognise an object as having agency and intentions and reason about its choices for deciding to act referred to as having intent recognition. Intent recognition can be considered a decision mechanism for forming understanding of multiple interdependent systems. People need to reason about and

* Corresponding author.

J. Zhou (Ed.): Complex 2009, Part II, LNICST 5, pp. 1631 – 1641, 2009.

solve the complex feedback loops that motivate behaviours of organisms whilst taking into account the state of the environment within which the judgment occurs.

Intentions are an agent's mental state of the purpose of their actions [1]. They pervade every aspect of our life as we create and execute plans to shape future events and satisfy needs and desires as they occur. At a high- level of analysis intentionality allows people with common goals to recognise their similarities and form society, cultures and groups. At a low-level intentionality allows participation in joint activities through the synchronisation of spontaneous planning and actions.

Despite giving an adaptive advantage in predicting the future behaviours of other systems (human or machine) intent recognition has weaknesses. Misunderstanding the meaning of others' actions can have consequences ranging in criticality when decisions of action must be made on those judgements. In favourable conditions such actions performed from a misinterpretation of intention may only result in simple embarrassment, while in critical scenarios it can result in the loss of life.

The increasing use of technology to distribute and perform tasks, between both spatial and temporal locations and between greater numbers of human components and computational components, are leading to the erosion of awareness of intent that makes joint working effective in collocated human-human systems. Judgments of intentionality by an observer are decided using the features, cues and attributes available from other actors (human or machine) and the environment. This allows other people's plans, goals and actions to be taken into account when deciding our own intentions, goals, plans and future actions thus adaptively avoiding dynamic conflicting problems. They also aid us in choosing when to pursue which goals, using the prediction of current and future enabled states. People make decisions about the world around them based on changes they perceive, interpret and understand from their own internal system (needs and desires) and those of external human and computational systems (behaviours). When examining groups of these systems interacting we must consider the effect that interpreted component behaviour has on components with agency to understand and predict the state of a system. The success or failure within these complex systems relies upon how we design, organise and manage the interactions between different human and computational components.

Currently when needing to make decisions about future events people lack the information support in human-computer systems to aid and enhance their judgments of intent, leaving them open to poor choices of action, conflict and failure. This oversight risks socio-technical systems heading towards costly and potentially harmful states.

This paper reports an exploratory study examining the phenomena of intentionality breakdowns to gain insight into their causes and occurrence. The following sections will describe previous relevant studies and theories concerning awareness, intention and breakdowns. Then the study methodology and results will be outlined before drawing conclusions.

2 Background

Today an imperative challenge in human-computer systems is how to design interactions between humans and increasingly independent autonomous systems (AS). These

systems are systems that "make and execute a decision to achieve a goal without full, direct human control" [2]. These types of systems are being deployed in a range of applications such as maritime, aerospace, medical healthcare and industrial [3]. The main benefit of an AS is that it provides an additional effort and capability for certain activities while reducing the need for human labour.

The evolution of command style computing to delegated autonomy means AS are now making an increasing number of decisions independently of human supervision and authorisation. People are becoming isolated observers trying to understand what is happening in a complex scenario and what is the most appropriate action required, in a situation where computers can choose to effect how a situation develops. An example are the supervisory operators of unmanned autonomous vehicles in defence environments [4]. This is the same problem people face in human-human interaction, namely coordinating effectively between multiple actors in complex scenarios. One avenue in which to pursue the human-AS problem is to understand how people cope with the complex uncertainty and unpredictability found in person-person interaction and joint activity. This could then be used as a basis for considering the less understood human-AS interaction requirements.

While the reviewed literature that follows provides an explanation of how intentions are recognised, this theory has not been tested or attempted to be applied in the design of human-computer systems. By considering how this might be applied to system design we can then go on to investigate how applicable it is and how well it enables awareness of intentions in human-AS interactions.

2.1 Awareness and Planning

By judging behaviours made by other actors in a scenario an observer gains insight into the state of components that would otherwise be costly to acquire. In other words they gain awareness – a better picture of how a system is evolving. Awareness has been defined by Endsley as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" [5]. They can also make causal attributions as to why an actor undertakes a particular course of action. Making such attributions can reveal both internal and external states of components and informs future knowledge of cause and effect, giving predictive powers.

People decide their future actions in a system through a combination of needs, awareness, goals and intentions. The decisions that they make and actions they perform can often be dependent on other actor's behaviour according to the degree of coupling of a given activity. For example, someone may choose their driven travel route to work according to known traffic conditions, journey duration, arrival time and an anticipation of the routes other drivers will take. Having information available to perception about the actions and states within a system and having the knowledge to understand it enables the generation of awareness of the effect on the system of a given action and therefore the means to choose the best course of action for the desired goal. In collocated settings intent recognition is aided by an actor having a greater perceptual awareness of the other agents and environment. But systems of remote distributed components require additional support to achieve even reduced levels of awareness and to enable coordinated behaviour. Currently when performing

shared and/or joint tasks that require interacting with people or technology in distributed environments systems provide support for a user's awareness of state and action. For example, a system allowing participants to monitor each other's tasks in a collaborative activity [6]. Hourizi & Johnson [7] have shown that this awareness can be significantly enhanced when information about future actions, intentions and implications is provided. In addition they have proposed and applied a framework to aid the design of technology to support this. Using this application of intentionality we hope to further explore this potentially beneficial solution.

2.2 Judgments of Intentionality

Psychologists have attributed the prediction of intentional components to the theory of mind. This cognitive mechanism recognises that motivations for our own behaviour can be different from that of other intentional components. Behaviours can be internally and externally generated which we can interpret in others as reasons or causes and make attributions about their motives. Intention has been defined as "a plan of action the organism chooses and commits itself to the pursuit of a goal - an intention thus includes both a means (action plan) as well as a goal" [8]. In dissecting this definition we can identify that actions have different types of purpose. Intentions can represent the end accomplishment of a plan and the tasks within the plan that enable the high level goal. This is recognising that predictions can be either short-term (the next action) or longer term (the end goal).

Considering the intention of an action as a prediction can serve two main benefits for an observer. From a hindsight perspective it can be used to understand and account for the reason for one particular action being performed over another i.e. where an action fitted into a plan after a goal has been achieved. In foresight understanding an intention can create a prediction and anticipation of the future possible states that other actors will try to create and how they will try to accomplish it. This is consistent with [9] who propose different types of intention – 'intention to' and an 'intention that' referring to an intention as an end goal or an intention as the purpose of an action enabling a state within a plan.

Judging intentionality is a decision-making process. The product of the process is dependent upon how the information available about a scenario is perceived and cognitively processed. Due to decisions of intentionality being dependent on the information available about an AS, false or lack of a critical level of information will result in an incorrect assumption about intentionality.

The research position taken here is that intentionality decisions in people are applied heuristics that cognitively "pattern match." Pattern matching uses observable information as cues to match to plans and actions of how to achieve goals in a given context stored as knowledge structures. Pattern matching occurs constantly as we observe a sequence of actions and construct an explanatory chain for each action to understand the end goal. Much of the motion and action processing occurs at an autonomic level [10], although we can consciously reason about intention using more effortful higher level reasoning. For observers this means a lack of information required for accurate matching, or incorrect information, will result in an incorrect match. The theoretical account adopted of how people are able to distinguish intentions is simulation theory [11]. As a synopsis the theory accounts for third-person mental

attributions by the attributor imaginatively seeing themselves from the targets perspective. They pretend to have the same initial states – such as desires and beliefs – and then simulate the decision being made. People do this by using the same mechanisms necessary to decide their own actions and running the target object's actor inputs into their mechanisms. The output from these mechanisms will then produce the target's intent.

We find the types of information inputted into such a mechanism in process models of intentionality such as Read and Miller's social dynamics model of person perception and dispositional inference [12]. They propose that inferred traits are frame-based structures or schemas where the goals of the agent being inferred play a central representational role. Read and Miller liken trait inference to the mind constructing "stories" or "narratives" where the trait is used to encode the perceived information as an exemplar to judge if the observed behaviour is a match. The encoding can include the actors, objects and actions perceived in an environment that are then assembled into a scenario. This scenario will activate in reasoning and memory possible goals for the action operating as a bottom-up and top-down process.

Studies by Malle and knobe [13] have found large amounts of consensus in the information used by people to determine if behaviours were intentional suggesting the mechanism uses a common strategy that can be supported through design. These factors included desire – what did they want; belief – how do they think they can get it; intention – did they decide to perform the action; awareness – do they know what they were doing; skill – are they able to turn intent into action.

2.3 Breakdowns

Activities involving coordinating tasks often occur between individuals with very different backgrounds and experience. It is this element of coordination that intent recognition plays a major role. Individual differences will often result in the decision to pursue different goals by different people in different ways, but in similar contexts. This means observers cannot rely on simplistic situational cause and effect to attribute reasons for behaviour. Individual intentions are chosen based on a range of factors that may be visible or hidden from an observer. For these reasons observers cannot be certain of correctly recognising the intention of an individual; however they can use their own expectations to assume a goal is being pursued until this is shown to be incorrect or fails. It is this use of expectations of what an observer thinks is occurring that causes breakdowns. Easterbrook defines a breakdown as "a mismatch between the internal mental models of people, such as the expectations of one participant in a collaborative activity and the actions of another participant in the same activity" [14]. The utility of examining breakdowns is twofold. Firstly, it provides a grounding tool through which to study intentionality recognition by revealing the requirements and limits of this mechanism. Secondly, it will allow the evaluation of systems for supporting intentionality recognition using these as a metric.

Related features of human cognitive failure similar to breakdowns and previously studied include errors, slips and mistakes investigated by James Reason [15]. However, these represent individual deviations of personal intention rather than observed intent.

3 Methodology

An exploratory study was conducted to capture a breadth of examples of intentionality breakdowns in naturalistic real-world settings. A problem with collecting data on these phenomena remains the internal, ephemeral nature of judgments and how to identify their occurrence. It is often difficult to identify a single causal factor in breakdowns. Reason says of the problem "we can only hope to indicate the commonly recurring patterns among these causative factors rather than attempt to make a precise determination of the origins of any one error" [16, p73]. Capturing these events that occur spontaneously is difficult when obtrusive research methods would affect the observer's perception of the system within which the judgment is made, thus changing its result.

Previous research methods for capturing breakdowns include video recording collaborative tasks. This allows a researcher to observe the video repeatedly to identify breakdowns according to their perceived deviations in the interaction [17]. Although we can try and judge breakdowns from inferences in behaviour deviation and reaction, only the actor in question has real firsthand knowledge of whether their expectations were incorrect, or if an observer misunderstood their intention. The study therefore used self-reporting diaries to capture examples. This was conducted by giving 4 participants notebooks where they recorded breakdowns as they went about their day-to-day life. Participants maintained the notebooks for 2 weeks. They were instructed to record a description of the breakdown scenario along with why they thought it happened. The research was interested in capturing errors where the participants had misinterpreted someone else's intention incorrectly, or someone else had interpreted their intention incorrectly.

4 Diary Results

The results provided 19 real life examples of people mistaking a person's intention in different ways. They raised a number of issues and guidance in the intentionality recognition needs to be supported. This section will report some key factors contributing to the breakdowns encountered. The purpose is too expose the limitations of intent recognition, and give the observer's attributed causes of these breakdowns thus demonstrating human decision-making dependencies.

4.1 Categorisation

Diary Extract 1

"The guy who takes tickets kept asking for my national rail card although it was already in his hand. Until he started looking through the tickets and picked up my oyster card. He looked at me and said this is for London. I said "I know" I keep it together for when I go there. After a short while I discovered what he meant by rail card. He meant the part of the ticket where the expiry date shows and it was behind the ticket..... I think the misunderstanding happened because he used the expression 'railcard' which meant to me the whole card rather than part of it."

The actors in extract 1 have a different notion of the artefact reference being used, that of the 'railcard' or train ticket. Often in our daily life interactions between cooperating people can involve asking for other actors to perform a task. People carry out different roles giving them different resources and responsibilities that we must negotiate with to satisfy our own goals. While in this example the observer believes they have met the request of the actor, the actor is unaware they have the required ticket in their hand and believe their goal conditions necessary for checking this are still disabled. When one actor directs another actor to perform an action, a lack of joint identification of the requested object will cause a breakdown. This is the result of people categorising objects in different ways.

4.2 Synchronisation

When coordinating actions another source of breakdowns was found to be a lack of feedback signals between actors. While some intentional behaviour is a signal in itself, action signals can be explicit or implicit. When an actor knows they are being observed they can change the action to aid recognition. But even when a goal is known, breakdowns can still occur because people's decision to act can be dependent upon knowing the time or sequence of when another actor will perform their behaviour if there is a necessary dependency.

Diary Extract 2

"Driving to Cornwall for the weekend reached the crossroads at the end of my road in Bath. It is busy so I stop. I need to go right and another car arrives at the other side of the crossroads. There are lots of cars coming along so neither of us can get out for a minute or so. Eventually a car wants to turn right into my road. She stops and gestures to me to come out. I check the traffic in the other direction and start to emerge. At the same time the car on the other side decides to come across also. We both notice and stop (otherwise we would collide)."

In extract 2 while both drivers knew the goal and plan of the other driver as indicated by car direction, there was no signal to coordinate when to perform actions within their plan (to pull away) to avoid the goals conflicting. Due to the length of time waiting at the crossroads and a degree of uncertainty whom had the right of way, both drivers acted when they recognised the goal conditions were enabled. They lacked the anticipation of when their actions would conflict, rather than if they would, and expected each other to give way for the other.

4.3 Expectations

In making judgments of intentionality people's explanation suggested they committed confirmation bias, similar to other decision processes [18]. This increased expectations and resulted in causes being recognised that were either not present, or had an equal chance of having a different cause. These could occur when the contextual situation matches a physical set of actions.

Diary Extract 3

"I was talking to my housemate. She was talking about the new exercise regime she had adopted to get fit. We were talking and I saw her go to the mirror. I thought this was to view her figure because she pulled her clothes tighter. When I said "I also pose in the mirror" she looked confused. She said "I felt I had something on me from the garden." Actually she thought there was some beetle or something on her because she had just come indoors from gardening."

Extract 3 demonstrates this priming effect through the conversation topic forming a context such that later actions were interpreted consistently through that context. The conversational context of exercise gave rise to expectations of how a mirror is used within that broader goal. It provided a plausible use for checking the progress towards the goal of getting fitter making that perceived behaviour explanation appear correct.

4.4 Pattern Matching

An unresolved visual problem with observing intention recognition is how to differentiate between near identical sets of actions that achieve different goals. Observers have particular difficulty with this kind of decision and this type of example made up the largest proportion of entries. Often it is a joint factor in breakdowns together with the presence of expectations.

Diary Extract 4

"Eating a meal in a restaurant with friends. Before the food came I was very hungry and thirsty. I had my bag with me and that had a bottle of water in it. I reached down to my bag under my chair without looking. I had to rummage around a bit in my bag before I could find it. By reaching down I had to bring my head closer to the table. There was a menu standing up on the table so my head was brought closer to it. My friend sitting opposite thought I was reading the menu and asked me what I was having."

In extract 4 the behaviour of leaning closer to the table top satisfied the action criteria for the goal of both reaching into the bag under the table and looking at the menu. The observer was unaware the bag was under the table so didn't know this goal was enabled and therefore as equally plausible as the menu goal.

5 Discussion

The diary study has successfully captured decision-making limitations in real-world scenarios. Despite the subjective nature of the data this approach contributes to the awareness needs of human decision-making. While the examples are not representative of all types of breakdown, they display features of human intent recognition warranting further investigation. The breakdown examples suggest Read and Miller's model of intent recognition [12] is a suitable abstraction for the process, but there are lower-level mechanisms needing further study. When deciding the intention of an actor a two-stage strategy is used to narrow the many possible goals. The first part is to generate expectations from a model of known goals and behaviours in the form of hypotheses to test, focusing perception on what to look for. Then these hypotheses

can be evaluated according to the goals that are known to be possible and plausible given the scenario conditions and the actor's likely motivation. Awareness or knowledge about these factors can change the decision of what intention is being performed.

When the human mind considers judgments about intentions it is considering many related factors in parallel and choosing the most likely possibility given the perceived and processed evidence. There will always be breakdowns that are unavoidable, since an action can be caused by more than one intention, certain behaviours will always display more than one goal. The only way to increase certainty is to supply more contextual or explicit intentional information, but this can become impractical. Even supplying all the correct information making people attend to it and understanding it, it will still result in different interpretations because people possess different knowledge of how to achieve goals and the way they process the same information can differ from day-to-day. The intent recognition mechanism has adapted not to prevent these mistakes, but to minimise them, given the time for behavioural responses.

This study generated a number of levels of information that can impact this decision by examining what information cues were missing or used. A key foundation of intention recognition is gaze or attention that identifies the target object over which an actor is making a decision. Studies have examined this mechanism [19], although the effect on intention recognition by supplying such information to observers requires examination. Another important feature to the production of possible goals of intent recognition is the similarity of the actor to the observer. If categorised as similar to the observer then consultation from their own mental states can occur since similar driving motives would produce appropriate goals. But if the actor is judged to be different then using prior knowledge such as stereotypes may be more appropriate. While less accurate this generalising technique can help to reduce the problem space considerably and any aid is preferable to none at all. Other information pertinent to the actor also includes knowing what is the relative importance of different goals to the actor. These may be internally or externally enforced. What role an actor plays in a scenario is one such external variable. Roles usually denote responsibility of tasks that will be prioritised over others in certain conditions. Other internal control can include the actor’s values, leading to certain beliefs about how to act in certain situations. At a high level this forms cultural trends and preferences. Boundaries of goal motivation are also restricted by physical states as well as mental states. The capability of the actor in ascertaining what goals they would realistically pursue means they are limited by their physical capability. The state of a system itself also enables some goals but prevents others. The physical location and state of components restrict actions and actors must enable some states before they can pursue other goals. It must be remembered that with such complexity to consider, the intent recognition mechanism was designed to provide a “good enough” answer by satisficing [20] to reconcile the conflict between decision time and accuracy. The computer support goal is therefore to influence the search for a plausible goal; to prevent its overrun when not enough information is supplied and to ease the process. Although both internal and external triggers cause breakdowns, technological design can primarily address the external triggers relating to the awareness of information at the levels of availability, perception and attention.

6 Conclusions and Future Work

Intent recognition is important for understanding how people deal with part of the complex decision-making within human-computer environments, and needs to be considered in systems featuring social components. We believe that supporting this mechanism holds promise for improving human decision-making in complex systems relating to attributing causality to components that possess agency. Supporting intent recognition to aid human-AS interaction requires the delivery, attendance and reflection of appropriate cues and context within which to process those cues. Providing the necessary features in technological design has potential benefits by enabling people to reduce the possible number of goals in a search space when making a judgment of intentionality. It is predicted this will reduce the number of opportunities for incorrect interpretation and resultant breakdowns or repairs. With further studies we aim to determine the effectiveness of supplying information based on current models of intentionality. Systems trying to support intentionality first need to address what information to provide and the effectiveness of doing so.

References

1. Malle, B.F., Moses, L.J., Baldwin, D.A. (eds.): Intentions and intentionality: Foundations of social cognition. The MIT Press, Cambridge (2001)
2. Connelly, J., Hong, W.S., Mahoney Jr., R.B., Sparrow, D.A.: Current challenges in autonomous vehicle development, Unmanned Systems Technology VIII. In: Proc. SPIE, 6230, 62300D (2006)
3. Watson, D.P., Scheidt, D.H.: Autonomous Systems. Johns Hopkins APL Technical Digest 26(4) (2005)
4. Liu, D., Wasson, R., Vincenzi, D.A.: Effects of System Automation Management Strategies and Multi-mission Operator-to-vehicle Ratio on Operator Performance in UAV Systems. Journal of Intelligent and Robotic Systems (forthcoming)
5. Endsley, M.R.: The application of human factors to the development of expert systems for advanced cockpits. In: Proc. of the Human Factors Society 31st Annual Meeting, Santa Monica, CA, pp. 1388–1392 (1988)
6. Kraut, R.E., Fussell, S.R., Siegel, J.: Visual information as a conversational resource in collaborative physical tasks. Human-Computer Interaction 18, 13–49 (2003)
7. Hourizi, R., Johnson, P.: Designing to support awareness: a predictive, composite model. In: Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI 2004, pp. 159–166. ACM Press, New York (2004)
8. Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H.: Understanding and sharing intentions: the origins of cultural cognition. Behav. Brain. Sci. 28, 675–691 (2005)
9. Panzarasa, P., Jennings, N.R., Norman, T.J.: Formalizing Collaborative Decision-Making and Practical Reasoning in Multi-Agent Systems. Journal of Logic and Computation 12(1), 55–117 (2002)
10. Scholl, B.J., Tremoulet, P.D.: Perceptual causality and animacy. Trends in Cognitive Science 4(8), 299–309 (2000)
11. Gordon, R.M.: Folk Psychology as Simulation. Mind and Language 1, 158–171 (1986)

12. Read, S.J., Miller, L.C.: Explanatory coherence and goal-based knowledge structures in making dispositional inferences. In: Malle, B., Hodges, S. (eds.) Other Minds: How humans bridge the divide between self and others, pp. 124–142. Guilford Press, New York (2005)
13. Malle, B.F., Knobe, J.: The folk concept of intentionality. Journal of Experimental Social Psychology 33(2), 101–121 (1997)
14. Easterbrook, S.M.: Coordination Breakdowns: Why Groupware is so Hard to Design. In: Proc. of 28th Hawaii International Conference on Systems Sciences (HICSS-28), Maui, Hawaii, USA (1995)
15. Reason, J.: Human Error. Cambridge University Press, Cambridge (1990)
16. Reason, J., Mycielska, K.: Absent-minded? The psychology of mental lapses and everyday errors. Prentice-Hall, Englewood Cliffs (1982)
17. Johnson, H., Hyde, J.: Towards modeling individual and collaborative construction of jigsaws using task knowledge structures (TKS). ACM Transactions on Computer-Human Interaction 10(4), 339–387 (2003)
18. Wason, P.C., Johnson-Laird, P.N.: Psychology of reasoning: Structure and content. D. T. Batsford, London (1972)
19. Phillips, A.T., Wellman, H.M., Spelke, E.S.: Infants' ability to connect gaze and emotional expression to intentional action. Cognition 85, 53–78 (2002)
20. Blythe, P.W., Todd, P.M., Miller, G.F.: How Motion Reveals Intention: Categorizing Social Interactions. In: Gigerenzer, G., Todd, P.M. (eds.) Simple Heuristics that Make Us Smart, pp. 257–285. Oxford University Press, Oxford (1999)