
Final accepted version (with author's formatting)

This version is available at: http://eprints.mdx.ac.uk/32174/

Copyright:

Middlesex University Research Repository makes the University's research available electronically.

Copyright and moral rights to this work are retained by the author and/or other copyright owners unless otherwise stated. The work is supplied on the understanding that any use for commercial gain is strictly forbidden. A copy may be downloaded for personal, non-commercial, research or study without prior permission and without charge.

Works, including theses and research projects, may not be reproduced in any format or medium, or extensive quotations taken from them, or their content changed in any way, without first obtaining permission in writing from the copyright holder(s). They may not be sold or exploited commercially in any format or medium without the prior written permission of the copyright holder(s).

Full bibliographic details must be given when referring to, or quoting from full items including the author’s name, the title of the work, publication details where relevant (place, publisher, date), pagination, and for theses or dissertations the awarding institution, the degree type awarded, and the date of the award.

If you believe that any material held in the repository infringes copyright law, please contact the Repository Team at Middlesex University via the following email address:
eprints@mdx.ac.uk

The item will be removed from the repository while any claim is being investigated.

See also repository copyright: re-use policy: http://eprints.mdx.ac.uk/policies.html#copy
Providing a foundation for interpretable autonomous agents through elicitation and modeling of criminal investigation pathways

Sam Hepenstal¹, Leishi Zhang², Neesha Kodogoda² and B.L. William Wong²
¹ Defence Science and Technology Laboratory, Porton Down, Salisbury, SP4 0JQ, UK
² Middlesex University London, The Burroughs, Hendon, London NW4 4BT, UK

Criminal investigations are guided by repetitive and time-consuming information retrieval tasks, often with high risk and high consequence. If Artificial intelligence (AI) systems can automate lines of inquiry, it could reduce the burden on analysts and allow them to focus their efforts on analysis. However, there is a critical need for algorithmic transparency to address ethical concerns. In this paper, we use data gathered from Cognitive Task Analysis (CTA) interviews of criminal intelligence analysts and perform a novel analysis method to elicit question networks. We show how these networks form an event tree, where events are consolidated by capturing analyst intentions. The event tree is simplified with a Dynamic Chain Event Graph (DCEG) that provides a foundation for transparent autonomous investigations.

Key Words
Autonomous agents; cognitive task analysis; XAI

INTRODUCTION

Criminal investigations involve high risk and high consequence situations in which it is vital that accurate and timely information is available to decision makers. Criminal intelligence analysts use this information for reasoning so that they, and their superiors, can make well informed decisions on important questions such as where to allocate resources, who to consider as potential suspects, or what risks a victim faces. Past research has considered the processes applied in criminal intelligence analysis, finding that they involve an “iterative combination of abductive, inductive and deductive inferences, information searching, associations, and further sense-making” (Wong and Kodagoda, 2016). Manual information searching comprises a significant proportion of this, where “each piece of insight leads to intense periods of manual information gathering” (Hepenstal et al. 2019b) and the initial investigation scope frames subsequent lines of inquiry. The potential benefits of speeding up this process are significant, where, if in a threat to life situation, reducing the time to find crucial information “could save someone’s life” (Hepenstal et al. 2019b). Artificial intelligence (AI) systems that can automate questioning to explore various investigation paths, therefore, present a significant opportunity for investigators to both speed up investigations and to challenge their initial scope. However, in such a high risk and high consequence domain there are important ethical concerns around bias and algorithmic opacity (Duquenoy, 2018) and these are potential barriers to the adoption of complex systems. A lack of understanding and oversight of algorithmic processes is identified as a serious issue by both system users, for example police officers (Babuta, 2019), and by human rights campaigners (Couchman, 2019). There are, therefore, critical design requirements for autonomous systems to be used in the context of criminal investigations, and a notable issue is the need for algorithmic transparency (Hepenstal et al. 2019a).

In this paper, we show the potential to model an event tree and a Dynamic Chain Event Graph (DCEG) (Barclay et al. 2013) that represents the lines of inquiry in an investigation. A DCEG is a discrete graphical model constructed from infinite event trees. We analyse data from Cognitive Task Analysis (CTA) (Klein et al. 1993) interviews with expert analysts, each interview covering a specific investigation scenario, to identify question networks and underlying intentions. We use these networks to form event trees that define a DCEG. We provide an example case to demonstrate how the DCEG identifies helpful investigation paths in a new investigation scenario. A DCEG simplifies infinite and complex option stages in an event tree, to form an accessible visual representation of the statistical model. A DCEG therefore provides a foundation to explore investigation paths, whilst clearly articulating them to analysts. In previous work, we have developed a conversational agent (CA) for information retrieval that provides algorithmic transparency of its reasoning (Hepenstal et al. 2020a). We used the Recognition-Primed Decision (RPD) model to deliver an explanation structure for intention concepts, in order to enhance user recognition and understanding of system behaviours. We propose to build upon this research to represent investigation pathways within a DCEG, where intention concepts inform the relationships between stages. We believe that this approach provides a platform for autonomous multi-stage reasoning, which tackles critical transparency issues for using AI systems in the field of criminal intelligence analysis.

CRIMINAL INVESTIGATIONS ARE HAMPERED BY INFORMATION OVERLOAD

In criminal investigations it is typical that analysts make repeated requests for information (Kodagoda and Wong 2016), with each new piece of insight requiring validation and triggering additional lines of inquiry. Much of this data processing is manual and time consuming, suffering from strict resource constraints. As explained by Mark Stokes, Head of Digital, Cyber and Communications Forensics Unit for the Metropolitan Police, “in digital forensics within England and Wales, the capacity to undertake what is required on criminal investigations is not there. We currently have a seven-month backlog.” (Stokes, 2018) Cressida Dick, Commissioner of the Metropolitan Police, reiterates the scale of information where she states, “there is so much data that has to be looked at…”
and “if police were able to harness data more effectively, a
‘very, very large proportion’ of crimes could be solved.”
(Shaw, 2019) Criminal investigations involve high risk and
high consequence scenarios and therefore the impacts of time
saving can be significant. Additionally, past research has
found that the scope of an investigation, while important to
help direct inquiries when resources are stretched, can also be
restrictive and introduce bias (Hepenstal et. al. 2019b).
Traditional analysis methods used to broaden analyst thinking
and address bias in investigations, such as analysis of
competing hypotheses (ACH) are also flawed (Dhami et. al.
2019). AI systems that can perform their own investigations
autonomously, whilst recommending information that may be
of interest to an analyst, have the potential to speed analysis
and challenge investigation scope without further burdening
analysts. This could include the identification of known and
unknown 'unknowns' (Logan, 2019). Even if a system can
explore only simple paths and make recommendations,
triggered by an initial question from an analyst, it could
provide helpful assistance.

SYSTEMS MUST EXPLAIN THEIR INFLUENCE

Systems are used across a wide range of domains to
make recommendations, for example to suggest items to buy
following an initial purchase or additional films to watch.
However, there are serious ethical considerations when it
comes to criminal intelligence analysis, where algorithmic
bias can have severe consequences. For example, if a system
directs investigation resources towards an innocent person,
through discriminatory processes.

Algorithmic bias can occur in various ways. “Human
error, prejudice, and misjudgement can enter into the
innovation lifecycle and create biases at any point in the
project delivery process from the preliminary stages of data
extraction, collection, and pre-processing to the critical phases
of problem formulation, model building, and implementation.”
(Leslie, 2019) Human rights campaigners have raised
concerns over the use of AI systems in the criminal justice
system, where “the nature of decision making by machines
means there is no option to challenge the process, or hold the
system to account.” (Couchman, 2019) Police analysts have
also raised concerns that an inability to understand and
challenge machine reasoning, and any bias that may have been
introduced, is a critical barrier to the use of complex systems
(Hepenstal et. al., 2019b). Central to the ability to challenge
and critique machine reasoning is the provision of algorithmic
transparency. In past work, we developed an algorithmic
transparency framework that identifies the need, in high risk
and high consequence domains, to provide both an explanation
of a system response together with the ability to inspect and
verify the goals and constraints of the system behaviour
(Hepenstal et. al., 2019a). Hoffman et. al. (2018) identify
some key concepts in literature on explaining systems, for
example, that explanation is a continuous process,
collaborative, triggered in specific situations, improves
learning and understanding, should clearly articulate caveats
and limitations, and should ensure the user understands what
is not being done as much as what is being done. Previously,
we have developed a transparent approach to interpret user
intentions when interacting with a conversational agent (CA)
(Hepenstal et. al., 2020). This approach can deliver
explanations that meet the key concepts. Crucially, we provide
an intention architecture that uses the way in which humans
recognise situations, the Recognition-Primed Decision (RPD)
model (Klein 1993, Hepenstal et. al. 2019b), to structure
explanations of the functional modules triggered by the
intention. This architecture allows a user to pick apart system
behaviours, in terms of the intention triggered by their input,
to clearly articulate caveats and limitations and to identify
contrasting intentions. The explanation structure has been
designed for a single stage interaction. The analyst asks a
question and triggers the CA to do some processing based
upon the matched intention. The analyst can step into the
answer, to see explanations of the various functional
processes, where intention attributes mirror the explanation
structure of the RPD model. For an AI system to be able to
conduct autonomous investigations it requires multiple stages
of processing. We propose that our intention architecture and
explanation structure aids us in providing transparency for
multi-stage processes, by combining multiple intentions in a
series.

ACCESSIBLE MODELLING OF EVENT SEQUENCES

Past approaches have looked to provide explainable
recommendations for event sequences. EventAction (Du et. al
2019), for example, can be used to recommend an action plan
to a student based upon their similarity to past students and
their desired outcome, such as to become a Professor.
EventAction models sequences of events as a probabilistic
suffix tree, based upon historic events, and applies a Markov
Decision Process (MDP) and Thompson Sampling to compute
and select a recommended action plan. A probabilistic event
tree could be a helpful way to explain possible inquiries at
each stage of an investigation, where with each response to a
question the analyst will have a set of options for how to
proceed. “Shafir demonstrated that an elicited tree was often
a much more powerful expression of an observer’s beliefs about
a process”, compared to other approaches to elicit a model
such as a Bayesian network (BN) (Shafir 1996, Smith and
Anderson 2008). Additionally, for capturing decision events in
an investigation, an event tree “provides a natural framework
through which time sequences can be incorporated” (Barclay
et. al., 2013). An event tree, however, can become complex to
represent visually as it grows. A Chain Event Graph (CEG)
can rectify this. “The CEG is derived from a probability tree
which is simplified into a CEG by introducing the concepts of
‘stages’ and ‘positions’. These group the vertices in the tree
together according to the associated conditional probabilities
on their edges.” (Barclay et. al., 2013) The graphical nature of
a CEG presents a useful opportunity for interpretability where
a user can see what variables have influence over others, and
can validate whether this is acceptable. Thwaites et. al. (2010)
find that, “as with Causal BNs, the identifiability of the effects
of causal manipulations when observations of the system are
incomplete can be verified simply by reference to the topology
of the CEG.” Chiappa and Isaac (2019), have demonstrated
that BNs are a “simple and intuitive visual tool for describing different possible unfairness scenarios underlying a dataset”, and this also applies to the CEG. CEGs therefore have useful qualities for providing transparency in criminal investigations, where it is important to trace back through reasoning steps and to understand what and how states, or questions, have influenced each subsequent piece of information gathered. The decisions made at each step seek to achieve some goal and communication of these and the underlying reasoning is crucial when developing observable autonomous systems. McDermott et. al. suggest that a system should understand the goals of the human users and communicate their intent in terms of what goals it is trying to accomplish for a task (McDermott et. al. 2018). A CEG that chains together cues, methods and goals for questions, is an effective foundation for observable autonomous reasoning.

Investigations involve repetitive questioning strategies. If, for example, a vehicle is presented in an output, an analyst may look for the owner and for any events in the database which have involved the vehicle. They may wish to do this every time a vehicle is found throughout their investigations. In these cases, the options available are repeated at different stages in the investigation and the event tree is infinite. We capture the topology of an infinite staged tree in a similar way to the CEG by using a Dynamic CEG (DCEG), as described by Barclay et. al. (2013). A DCEG provides a succinct explanation of the stages available and their influences on one another, which could help to achieve algorithmic transparency. In our system event selections trigger complex processes that themselves require explaining, and we utilise our intention architecture (Hepenstal et. al., 2019b, Hepenstal et. al., 2020a), underpinned with an explanation structure reflecting the RPD model. In this paper, we model an event tree to form a DCEG (Barclay et. al., 2013). The DCEG is a useful aid to explore possible investigation paths, where each state reflects the explanation structure of the relevant intention.

### Table 1: Question Elicitation Example

<table>
<thead>
<tr>
<th>Statement [CTA; Analyst 4; 2.00 -&gt; 10.00] (Input=Suspect Phone Number)</th>
<th>Specific Need</th>
<th>Specific Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;3 things you do instantly with the number. Stick it through your (databases), see if any other existing links. ... Check with all call data we have collected from operation... run subscriber checks on numbers he has called to get info on contract subscriber... We then go and find other phone calls (involving suspect phone number)... Can get call data for others in the network... Also check all numbers additional people have phoned against all other numbers (in databases).&quot;</td>
<td>See if any other existing links for phone number</td>
<td>Is the phone number connected to known events?</td>
</tr>
<tr>
<td></td>
<td>Check with all call data and subscriber checks</td>
<td>What people are associated to phone number?</td>
</tr>
<tr>
<td></td>
<td>Find other phone calls</td>
<td>What calls involved this phone number?</td>
</tr>
<tr>
<td></td>
<td>Get call data for others in network</td>
<td>What numbers are being called?</td>
</tr>
<tr>
<td></td>
<td>Check all numbers additional people have phoned against all other numbers</td>
<td>What calls involved these numbers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>What known events are connected to numbers?</td>
</tr>
</tbody>
</table>

### INTERVIEW DATA

We conducted Cognitive Task Analysis (CTA) interviews, applying the Critical Decision Method (CDM), with four criminal intelligence analysts. In each interview we delved into a particularly memorable investigation that they were involved in from start to end. For this study, we were most interested in how analysts questioned data as they sought to retrieve information to advance their investigations, in particular how questions led to insights that triggered subsequent inquiries. Each analyst had more than 3 years’ experience. We have previously analysed the interview data to identify distinct questions and to structure them against the Recognition-Primed Decision (RPD) model (Hepenstal et. al., 2019b, IUI ExSS 2020). In this study we revisit these questions and identify links between them, where a link can be drawn if the result of one question is subsequently used to form cues in another. The questions could be asked of a conversational agent (CA) to retrieve the information required and we propose that question networks can be captured dynamically in the future through such interactions.

**Figure 1:** Question Network for Firearm Scenario

### ANALYSIS: ELICITING QUESTION NETWORKS FROM INTERVIEW DATA

Drawing upon timeline analysis of the investigations (Hepenstal et. al., 2019b), we extracted questions asked by analysts in the order in which they were considered. Analyst questions were related to specific input cues, for example an entity found in the results of a previous inquiry. Interview statements tend to describe general processes performed in the investigation; Table 1 presents some examples to demonstrate how questions were elicited. We have attempted to capture the underlying information needs from each statement, and have extracted from this specific questions that could address requirements. Analyst questions are not asked in isolation, where there are directed relationships if the outputs of one feed another. The relationships between questions can be more clearly presented as a network and we applied a novel analysis technique to form networks from the interview data. Figure 1 shows a network of questions for one of the interview scenarios.
forms an event tree for questioning: example case results

We have used our previous work on developing intention concepts (Hepenstal et al., 2019b, Hepenstal et al., 2020) to consolidate question events for two interview scenarios, a kidnapping and a firearm dealing, allowing us to build an abstract question network. In our abstract network a node (question event) requires three components: an input i.e. the question subject (e.g. a phone number), a query class (e.g. people), and an intention. The intention defines the way in which the question will be processed. We can make our event stages more or less domain specific by manipulating class granularity.

A third interview scenario involved an attempted murder, where the female victim was found alone in her house and the incident was reported by her husband, who claimed to be on the phone with her at the time of the assault. The analyst explained that initially the husband was a suspect, due to ongoing divorce proceedings between himself and the victim. At the outset of the investigation, the analyst was responsible for verifying the husband’s statement, “to identify conflicts” [A4; 25:00], where the key piece of information available was the husband’s phone number. In Figure 2, we show an event tree for this scenario where vertex stages have the prefix ‘v’.

Despite situational differences in the scenario, by identifying a starting stage and input we can build a tree that describes all the questions the analyst asked. The network of investigation questions shown in Figure 3 are represented within event tree stages (Figure 2) and roughly reflects analyst questioning, where they explain that “(the) husband alerted officers, he said she had been on the phone to him when the burglary and attack occurred. I looked at telephone records to confirm that his statement is correct, backed up by phone records and corroborating information. The analyst is trying to back up or refute information in the statement. I’m asking for phone numbers within (call) data which are matched and person details, such as those involved in scene. I look at the network of which phones have called each other.” [CTA; A2: 24:00] This example demonstrates the potential to predict plausible question networks, even for new scenarios.

Outcome: a model for investigation questioning

Even in our simple example, there are 19 distinct stages and the tree is infinite. To represent the entirety of the tree is therefore not possible, yet the current tree does not clearly reflect the influences between states. A DCEG representation can be used to represent the tree “in a much more compact and easily interpretable form.” (Barclay et al. 2013) To build the DCEG, we identified probabilistic symmetries in the tree, i.e. stages with identical probabilities across options available. In our example, these are the revisited stages. Representing the infinite event tree as a DCEG reduces the complexity significantly to 7 positions. The topology of the DCEG allows us to assess influences across positions, where we can inspect and verify information held at each of the cues and intentions that define relationships between them. The relationships in our DCEG introduce their own complexities and related methods for information retrieval. In our DCEG a possible path, for example, is to find equipment involved in event data, then find additional events that are linked to the equipment, before exploring these events further. However, a relationship does not exist to find events connected to the equipment and explore these further. Instead, connected events lead to the end position, as these are identified as a goal based upon past investigations. The possible paths are constrained and an analyst must be aware of this when interpreting results. We can use our intention architecture and explanation structure that reflects the RPD model to enhance analyst recognition of the goals and constraints of selected paths.
FUTURE WORK: A FOUNDATION FOR AUTONOMOUS QUESTIONING

Under certain assumptions, a DCEG “corresponds directly to a semi-Markov process.” (Barclay et. al., 2013) Therefore, in a similar fashion to EventAction, it is possible to generate and select interesting lines of inquiry for an investigation from our DCEG. The automation of information retrieval could be valuable to analysts, who spend “quite a lot (of time) doing detective work, where a piece of intelligence is nothing on its own, but we needed to trawl data and find links.” [CTA; A4: 2.30] In the kidnapping scenario it took 16 question stages to gather the information required to identify where the person had been taken. Likewise, it took 12 question stages in the firearms dealing scenario to address their "need to find this person (firearm dealer).” [CTA; A4; 09:00] An autonomous approach would save much human effort. However, an analyst needs to sufficiently inspect and verify the reasoning, goals and constraints of the methods applied. We provide a foundation for observable automation.

CONCLUSION

We have demonstrated a novel approach to elicit question networks from interview data. In future work, we will look to capture these dynamically from interactions with a CA. The question networks form an event tree and a DCEG, from which we can generate and select lines of inquiry with an explanation structure at their foundation. Further consideration is needed for how to define goals in investigation paths. We propose that we can capture a better understanding of analyst processes and we should consider how to utilise this. Our approach could also develop transparent autonomous aids for other high risk and high consequence domains, such as medical diagnoses.

ACKNOWLEDGEMENTS

This research was assisted by experienced intelligence analysts from the Defence Science Technology Laboratory (Dstl).

REFERENCES