Erotetic Search Scenarios: Revealing Interrogator's Hidden Agenda

Mariusz Urbański
Chair of Logic and Cognitive Science
Institute of Psychology
Adam Mickiewicz University
Poznań, Poland
Mariusz.Urbanski@amu.edu.pl

Paweł Łupkowski
Chair of Logic and Cognitive Science
Institute of Psychology
Adam Mickiewicz University
Poznań, Poland
Pawel.Lupkowski@amu.edu.pl

Abstract

The way of formal modeling of a hidden agenda of an interrogator is described in terms of Inferential Erotetic Logic. Two examples are given: one is based on a simple detective story, the other is based on an analysis of a judge's strategy in the Turing Test.

1 Introduction: erotetic basis

Wiśniewski (2003) defines erotetic search scenarios (e-scenarios for short) within the framework of Inferential Erotetic Logic (IEL) as a possible technique for solving problems expressed by questions. He claims that:

> When a problem is expressed by a question which has a well-defined set of direct[1] [...] answers, one can [...] apply an e-scenario in order to find the solution to the problem. Viewed pragmatically, an e-scenario provides us with conditional instructions which tell us what questions should be asked and when they should be asked. Moreover, an e-scenario shows where to go if such-and-such a direct answer to a query appears to be acceptable and does so with respect to any direct answer to each query (Wiśniewski, 2003, p. 422).

Thus an e-scenario may be interpreted as a plan for an interrogation (the questioned being a human, a database, an Oracle etc.) that describes a "hidden agenda" of an interrogator.

Suppose that I am questioning a certain suspect in order to determine if this person is guilty or not (for definiteness let my problem be expressed by a question: "Who stole the tarts?"). In such a situation addressing the question directly may not be the most brilliant idea, unless the suspect is willing to plead guilty. If I am interested in something more than a declaration of a person in question I have to seek for a more or less indirect solution by gathering evidence and by making inferences on its basis. Description of this evidence and a plan for further inferences forms in this case my hidden agenda. It expounds *a*) initial information relevant to the case and *b*) inferential steps made on its basis. Inferential steps involved are possibly of two kinds: standard declarative ones and erotetic ones. Erotetic inferences are these in which questions play the role of conclusion and/or premises. Questions arise when there is a gap in available information (initial or derived), but from the investigator's point of view it is important to pose only such auxiliary questions that are both informative and cognitively useful, that is, answers to which are helpful in answering the initial question. This may be formally explicated in terms of erotetic implication, an erotetic counterpart to the entailment relation (Wiśniewski, 1995):[2]

Definition 1 *A question* Q implies *a question* Q^* on the basis of *a set of d-wffs* X *(in symbols:*

[1] *Direct answers* are the answers which "are directly and precisely responsive to the question, giving neither more nor less information than what is called for" (Belnap, 1969, p. 124). For the sake of generality they may be called *principal possible answers* (Wiśniewski and Pogonowski, 2010).

[2] Our language is the language of First-order Logic enriched with question-forming operator ? and brackets $\{,\}$ (call this language $\mathbf{L}$). Well formed formulas of FoL (defined as usual; additionally, we allow for names of formulas to appear as arguments of predicate symbols) are *declarative well-formed formulas* of $\mathbf{L}$ (d-wffs for short). Expressions of the form $?\{A_1, \ldots, A_n\}$ are *questions* or *erotetic formulas* of $\mathbf{L}$ (e-formulas for short) provided that $A_1, \ldots, A_n$ are syntactically distinct d-wffs and that $n > 1$. The set $\mathbf{d}Q = \{A_1, \ldots, A_n\}$ is the set of all the direct answers to the question $Q = ?\{A_1, \ldots, A_n\}$. Thus an erotetic formula $?\{A, \neg A\}$ expresses a simple yes-no question: "Is it the case that A or is it the case that $\neg A$?"; this kind of questions we shall abbreviate by $?A$. Let also the symbol T stand for any logically valid formula. Intuitively, T stands for the lack of factual knowledge: a question of the form $?\{A_1, \ldots, A_n, \mathsf{T}\}$ reads "Is it the case that A_1 or ... or is it the case that A_n or no required information is available?" (Wiśniewski, 2007).

$Im(Q, X, Q^*)$) iff

1. *for each direct answer A to the question Q: $X \cup \{A\}$ entails the disjunction of all the direct answers to the question Q^*, and*

2. *for each direct answer B to the question Q^* there exists a non-empty proper subset Y of the set of direct answers to the question Q such that $X \cup \{B\}$ entails the disjunction of all the elements of Y.*

If $X = \emptyset$, then we say that Q implies Q^* and we write $Im(Q, Q^*)$.

The first condition requires that if the implying question is sound[3] and all the declarative premises are true, then the implied question is sound as well[4]. The second condition requires that each answer to the implied question is potentially useful, on the basis of declarative premises, for finding an answer to the implying question. To put it informally: each answer to the implied question Q^*, on the basis of X, narrows down the set of plausible answers to the implying question Q.

Consider a simple example. My initial question is;

(Q) Who stole the tarts?

Suppose that I manage to establish the following evidence:

(E_1) It is one of the courtiers of the Queen of Hearts attending the afternoon tea-party who stole the tarts.

Thus my initial question together with the evidence implies the question:

(Q^*) Which of the Queen of Hearts' courtiers attended the afternoon tea-party?

If moreover I know that:

(E_2) Queen of Hearts invites for a tea-party only these courtiers who made her laughing the previous day.

then Q^* and E_2 imply the question:

(Q^{**}) Which courtiers made the Queen of Hearts laughing the previous day?

[3]A question Q is *sound* iff it has a true direct answer (with respect to the underlying semantics).

[4]This property may be conceived as an analogue to the truth-preservation property of deductive schemes of inference.

Erotetic search scenarios may be defined as sets of so-called erotetic derivations (Wiśniewski, 2003) or, in a more straightforward way, as finite trees (Wiśniewski, 2010, p. 27–29):

Definition 2 *An* e-scenario *for a question Q relative to a set of d-wffs X is a finite tree Φ such that:*

1. *the nodes of Φ are (occurrences of) questions and d-wffs; they are called e-nodes and d-nodes, respectively;*

2. *Q is the root of Φ;*

3. *each leaf of Φ is a direct answer to Q;*

4. *$\mathbf{d}Q \cap X = \emptyset$;*

5. *each d-node of Φ:*
   (a) *is an element of X, or*
   (b) *is a direct answer to an e-node of Φ different from the root Q, or*
   (c) *is entailed by (a set of) d-nodes which precede the d-node in Φ;*

6. *for each e-node Q^* of Φ different from the root Q:*
   (a) *$\mathbf{d}Q^* \neq \mathbf{d}Q$ and*
   (b) *$Im(Q^{**}, Q^*)$ for some e-node Q^{**} of Φ which precedes Q^* in Φ, or*
   (c) *$Im(Q^{**}, \{A_1, ..., A_n\}, Q^*)$ for some e-node Q^{**} and some d-nodes $A_1, ..., A_n$ of Φ that precede Q^* in Φ;*

7. *each d-node has at most one immediate successor;*

8. *an immediate successor of an e-node different from the root Q is either a direct answer to the e-node, or exactly one e-node;*

9. *if the immediate successor of an e-node Q^* is not an e-node, then each direct answer to Q^* is an immediate successor of Q^*.*

Definition 3 *A* query *of an e-scenario Φ is an e-node Q^* of Φ different from the root of Φ and such that the immediate successors of Q^* are the direct answers to Q^*.*

We shall elaborate the idea of representing interrogator's hidden agenda via e-scenarios on two examples. The first one is a simple detective story based on a Smullyan's (1978) logical puzzle. The second one is an analysis of a judge's strategy in the Turing Test.

2 Vampires, zombies and humans

On a certain island the inhabitants have been bewitched by some kind of magic. Half of them turned into zombies, the other half turned into vampires. The zombies and the vampires of this island do not behave like the conventional ones (if any): the zombies move about and talk in as lively a fashion as do the humans, and the vampires even prefer drinking strong mocca over anything else. It's just that the zombies of this island always lie and the vampires of this island always tell the truth[5]. What is also important, both vampires and zombies never miss a reasonable opportunity to tell the truth or to lie, respectively. Thus they always do their best to answer questions addressed to them.

A native named Eugene has been suspected of an attempt to break in an ATM near the police station. The case has been assigned to Inspector Negombo (a vampire) of local police force. His first task was to establish if the accused is a vampire or a zombie. Inspector Negombo was clever enough to determine that Eugene is a vampire on the basis of the suspected's answer to a single question. What was Negombo's question?

There are many possibilities. The question could be *e.g.* "Is it the case that you a vampire or you are not a vampire?": the positive answer identifies the answerer as a vampire, the negative one identifies the person as a zombie. This solution may cause a usual astonishment of the Watson-like audience as well as rise the admiration of Negombo's methods[6]. However, its explanation would certainly cause as usual "It is pretty obvious now" reaction.

Let us reveal Inspector Negombo's hidden agenda. He knows the following fact:

1. Every native is either a vampire or a zombie.

He knows also the following rules:

2. Every native who is a vampire utters true sentences.

3. Every native who utters a true sentence is a vampire.

(These two rules could be expressed more succintly as an equivalence, but it is in our hidden agenda to leave them in an implicational form.) The only problem is to operationalise the rules in such a way that they will be applicable to questions. This may be done as follows:

$2'$. For every native x, if x gives back a true answer to a posed question, then x is a vampire.

$3'$. For every native x, if x is a vampire, then x gives back a true answer to a posed question.

An important premise in Negombo's reasoning is the following:

4. Eugene is a native of the island.

What remains is to find a suitable question. It would be useless to ask questions like: "Are you a vampire?". Every native of the island would answer this questions positively giving no clue who is lying and who is telling the truth. The point is to ask a question with such direct answers that both Negombo and the suspected will know their truth values – as in the case of questions about fairly simple logically true (or false) sentences. Let us express Negombo's hidden agenda in terms of a formalized language (Wiśniewski, 1995).

Let $V(x)$, $Z(x)$, $N(x)$ stand for expressions: "x is a vampire", "x is a zombie", "x is a native of the island" respectively, and let $U(x, A_i, ?\{A_1, \ldots, A_n\})$ stand for an expression "x gives back an answer A_i to the question $?\{A_1, \ldots, A_n\}$"[7] (provided that $i = 1, \ldots, n$). Finally, let the constant a represent Eugene. Negombo's agenda is depicted by the erotetic search scenario of example 1 (for brevity we assume that R stands for the formula $V(a) \vee \neg V(a)$).

Negombo's initial question, "Is Eugene a vampire, a zombie, or no information is available?" is expressed by the first e-formula: $?\{V(a), Z(a), \top\}$. The inspector makes use of four declarative premises. The first one, $\forall x(N(x) \rightarrow V(x) \perp Zx)$, expresses the fact 1, that every native is either a vampire or a zombie. The second and the third premises express the rules $2'$ and $3'$, respectively. The fourth premise states that Eugene is a native of the island.

[5] An important though tacit assumption is that both vampires and zombies of this island reason in accordance with classical logic.

[6] Although if a reader would like to argue that on this particular island this should be something like a standard procedure of interrogation, we shall agree.

[7] In a formula $U(x, A_i, ?\{A_1, \ldots, A_n\})$ the second argument of the predicate symbol U is a name of a d-wff and the third argument is a name of an e-formula. For the sake of brevity we omit the quotation marks, as no ambiguity arises in this context.

Example 1

$$?\{V(a), Z(a), \mathsf{T}\}$$
$$\forall x(N(x) \to V(x) \perp Z(x))$$
$$\forall x(U(x, R, ?R) \to V(x))$$
$$\forall x(V(x) \to U(x, R, ?R))$$
$$N(a)$$
$$?V(a)$$
$$?U(a, R, ?R)$$

$$\begin{array}{cc} U(a, R, ?R) & \neg U(a, R, ?R) \\ V(a) & \neg V(a) \\ & Z(a) \end{array}$$

The initial question together with the first and the fourth declarative premise implies the question $?V(a)$[8]. Observe that this question, though implied, is not asked: it is not a query of the scenario[9]. The question $?V(a)$ together with the rules implies in turn the next question in the scenario, expressed by the erotetic formula $?U(a, R, ?R)$. This question is a query: it is answered in the scenario. The positive answer to it, expressed by the formula $U(a, R, ?R)$, together with the rule $2'$ entails the positive answer to $?V(a)$ which is also an answer to the initial question. On the other hand, the negative answer, expressed by the formula $\neg U(a, R, ?R)$, together with the rule $3'$ entails the negative answer to $?V(a)$ which, in turn, together with the first and the fourth premise, entails $Z(a)$, an answer to the initial question[10].

What this scenario shows is that the only question that needs to be actually addressed to Eugene is the question $?R$. What is more, this question and Eugene's answer would form all the interrogation conducted by Negombo. However, $?R$ is only mentioned and not used in the scenario (it is neither the initial question nor an implied question). The scenario reveals also what conclusions will Negombo derive on the basis of his declarative premises, questions and possible Eugene's answers[11].

One important property of e-scenarios is that they are "conditionally safe": if an initial question Q of a given e-scenario Φ has a true direct answer and if all the declarative premises of Φ are true, then at least one path of Φ leads to a true answer to Q; what is more, all d-wffs of this path are true and all e-formulas of this path have true direct answers as well. This is the essence of the Golden Path Theorem (Wiśniewski, 2003, p. 410–411).

It would be quite simple now to solve the problem: just ask Eugene if he really did try to break in the ATM. But this simple plan has been ruined by discovery that one premise on which Negombo's inferences were dependent is false: Eugene was not a native of the island. He came there as an immigrant from the nearby island, inhabited exclusively by humans, who are totally unpredictable as for the truth or falsity of what they tell[12]. Moreover, Eugene refused to give any sort of testimony. A short investigation among Eugene's friends (all confirmed being vampires by Negombo's test) led Negombo to establish the following rules:

5. If Eugene did run short of money, then he attempted to break in an ATM or borrowed some money from Eustace.
6. If Eugene didn't run short of money, then he went shopping or visited his favourite pub.
7. If Eugene attempted to break in an ATM or went shopping, then he has been seen in a local mall.
8. If Eugene borrowed some money from Eustace or visited his favourite bar, then he hasn't been seen in a local mall.

On this basis Inspector Negombo devised the plan for further interrogation of Elyssa, the only of Eugene's friends able to describe the course of events of that particular day (cf. example 2; $At(x)$ stands for "x attempted to break in an ATM", $Sh(x)$ stands for "x ran short of money", $Br(x, y)$ stands for "x borrowed money from y", $Sp(x)$ stands for "x went shopping", $Pb(x)$ stands for "x visited his favourite pub", $Ml(x)$ stands for "x

[8] In fact the question $?V(a)$ is implied by the question $?\{V(a), Z(a), \mathsf{T}\}$ on the basis of the empty set as well, but this holds for trivial reasons. Questions with T as a direct answer should be dealt with with some caution if one is not to fall into such triviality. One way to provide such caution is to employ *constructive* erotetic implication, instead of the standard one (Wiśniewski, 2007).

[9] However, this question is an important inferential step, because erotetic implication is not transitive.

[10] The relevant erotetic implications on which this and subsequent scenarios are based may be found in (Wiśniewski, 1995), (Wiśniewski, 2007), and (Łupkowski, 2010).

[11] Van Kuppevelt (1995) presents somewhat similar ideas of deriving questions (both explicit and implicit) from other questions and/or declaratives in the context of analysis of discourse structure. However, his informal account concerns wh-questions only.

[12] Although still *classically* unpredictable.

Example 2

$$?\{At(a), \neg At(a), \mathsf{T}\}$$
$$Sh(a) \rightarrow At(a) \vee Br(a,c)$$
$$\neg Sh(a) \rightarrow Sp(a) \vee Pb(a)$$
$$At(a) \vee Sp(a) \rightarrow Ml(a)$$
$$Br(a,c) \vee Pb(a) \rightarrow \neg Ml(a)$$
$$?Ml(a)$$

$$Ml(a) \qquad\qquad \neg Ml(a)$$
$$\neg(Br(a,c) \vee Pb(a)) \qquad\qquad \neg(At(a) \vee Sp(a))$$
$$\neg Br(a,c) \qquad\qquad \neg At(a)$$
$$\neg Pb(a)$$
$$?Sh(a)$$

$$Sh(a) \qquad\qquad \neg Sh(a)$$
$$At(a) \vee Br(a,c) \qquad\qquad Sp(a) \vee Pb(a)$$
$$At(a) \qquad\qquad Sp(a)$$
$$\mathsf{T}$$

Example 3

$$?L(b, \neg Ml(a))$$
$$U(b, \neg Ml(a), ?Ml(a))$$
$$\forall x(S(x, B_1, A) \rightarrow L(x, A))$$
$$\ldots$$
$$\forall x(S(x, B_n, A) \rightarrow L(x, A))$$
$$\forall x(\neg S(x, B_1, A) \wedge \ldots \wedge \neg S(x, B_n, A) \rightarrow \neg L(x, A))$$
$$?\{L(b, \neg Ml(a)), \neg L(b, \neg Ml(a)), S(b, B_1, \neg Ml(a))\}$$
$$?S(b, B_1, \neg Ml(a))$$

$$S(b, B_1, \neg Ml(a)) \qquad \neg S(b, B_1, \neg Ml(a))$$
$$L(b, \neg Ml(a)) \qquad \ldots$$
$$?S(b, B_n, \neg Ml(a))$$

$$S(b, B_n, \neg Ml(a)) \qquad \neg S(b, B_n, \neg Ml(a))$$
$$L(b, \neg Ml(a)) \qquad \neg L(b, \neg Ml(a))$$

has been seen in a local mall" and c stands for Eustace).

"If he's guilty and if this is my lucky day, I'll send him to the court in two questions", Negombo said to himself. "If he's been in the mall but didn't run short of money then my information is insufficient and I will need new evidence. If he hasn't been in the mall, he's innocent. Well, we'll see". He ordered one of his lieutnants to conduct an interrogation according to this plan[13]. The lieutnant soon reported the outcome: Elyssa answered first query with "No". Eugene hasn't been in the mall. He's innocent!

However, it occured that Elyssa is a human, too. Unfortunately for Elyssa and Eugene, Negombo studied nonverbal behaviour of human liars (Ekman, 2001) and has been identified as a "Truth Wizard", that is, a person who can identify deception with exceptional accuracy of more than 80% (Harrington, 2009). Besides his natural talent he devised for himself a list of behaviours that help identifying lies with satisfactory precision. Negombo, highly suspicious as for the truth of what Elyssa testified, decided to try to kill two birds with one stone and possibly accuse her of false testimony. Negombo repeated Elyssa's interrogation, this time personally, having in mind an agenda represented by the e-scenario of example 3 ($L(x, A)$ stands for "x lies saying A", $S(x, B, A)$ stands for "x expresses the set of behaviours B while saying A" and b stands for Elyssa).

[13]The reader may notice that both queries of this scenario, that is $?Ml(a)$ and $?Sh(a)$, might demand plans for investigation in the form of e-scenarios on themselves. Such auxiliary e-scenarios can be incorporated into the main one by the embedding operation (Wiśniewski, 2003).

Again, this scenario shows that the only question that Negombo should actually pose to Elyssa is "Has Eugene been in the mall?" (the one represented by the e-formula $?Ml(a)$) although it is known what will the answer be. All the remaining questions play the role of milestones on Negombo's way of thinking in solving the initial problem. Notice that they are concerned not with the content of Elyssa statement but with the way she provided that statement.

Elyssa repeated her previous testimony that Eugene has not been in the mall. But saying this she expressed a set of behaviours characteristic for a liar (say that they were microexpressions of her lips indicating disbelief in what she has been saying[14]). On this basis Negombo determined that she is lying that Eugene has not been in the mall. To finish his investigation quickly he decided to employ ethically disputable means. He produced a fake witness (who testified that he has seen Eugene in the mall) and confronted Elyssa with him. Elyssa finally admitted that she was lying and that Eugene in fact has been in the mall on that particular day. Her statement has been recorded. The case has been sent to the court.

3 The Turing Test

The idea of erotetic reconstruction of an interrogator's hidden agenda comes from Łupkowski's (2010) analysis of a judge's strategies in the Turing Test (TT).

[14]An interested reader may have a look at the example of such microexpression on Paul Ekman's page: `http://www.paulekman.com`.

In the TT the judge (interrogator, **J**) poses questions and his/her aim is to establish on the basis of answers received if an interlocutor (answerer, **A**) is a human or not (in which case it is inferred that the interlocutor is a machine). It is reasonable to assume that the judge would not ask questions at random, but will use some kind of a strategy. This is due to the fact that it is the judge who is responsible for the way the test is performed – **J** asks questions and **A** answers them. It is **J** who also decides when the test will end. There are good reasons why it is beneficial for **J** to choose an e-scenario as a questioning plan in the TT:

- An e-scenario gives information when and what question should be asked (relative to the initial question and initial premises).
- What is more, it ensures that all subsequent questions asked would be relevant to the initial question (so we may say that no unnecessary information would be collected by **J**).
- E-scenarios also guarantee that each subsequent question asked is a step towards the answer to the initial question.
- The Golden Path Theorem guarantees that for a strategy expressed by an e-scenario there exist at least one such path that ends with the answer to the initial question which is true (relative to the initial premises).
- A strategy presented by an e-scenario is flexible in the sense that it can be modified and rearranged by the embedding procedure to fit the **J**'s needs for the current interrogation.

It would be useless for **J** to ask "Are you a human?", because everything he/she can get are just simply **A**'s doubtful declarations. A sound strategy for **J** is to devise a plan for an interrogation that is based on assumptions concerning criteria for being a human. The judge may formulate this criteria as sufficient and/or necessary conditions, representing his/her expectations as for the answers which would be given to his/her questions by a human.

One of the possible ways to put the judge's beliefs might be the following:

"If d is a human, then d fulfils a condition C_i."

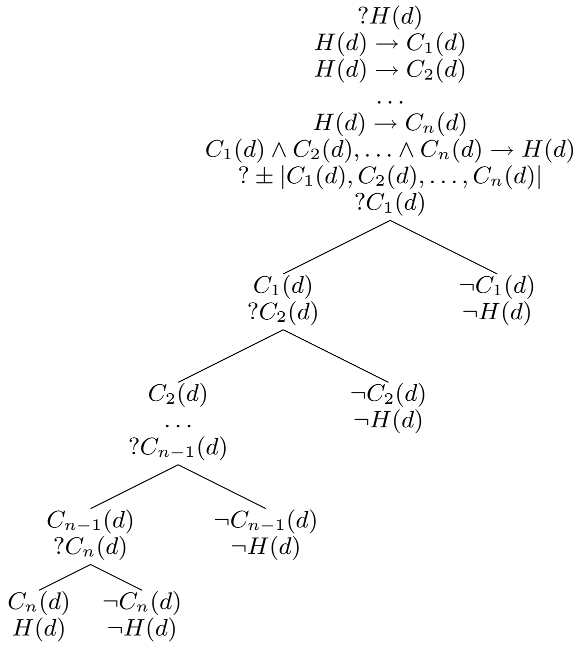
We may present the interrogator's beliefs as a set of formulas of the following form:

$$H(d) \to C_1(d)$$
$$H(d) \to C_2(d)$$
$$\dots$$
$$H(d) \to C_n(d)$$
$$C_1(d) \land C_2(d) \land \dots \land C_n(d) \to H(d)$$

where H (standing for "... is a human") is different from any of C_i ($1 \leq i \leq n$).

The interrogator uses the necessary conditions of being a human in this case. Fulfilling all these rules together is – in the interrogator's opinion – a sufficient condition of being a human. The e-scenario which might be used as a strategy for the interrogator in this case is presented as example 4.[15]

Example 4



The queries of this e-scenario should be treated as questions asked by **J** to himself/herself. Again, it seems fruitless to ask these questions directly to **A**. If **J** for example asks a question "Can you play chess?" and **A** answers "Yes", **J** do not obtain any interesting piece of information (at least if we consider the Turing Test's perspective). This is the reason why we will treat the queries of an e-scenario as only setting out the questions which should be asked to the answerer.

[15] In this scenario we make use of a *conjunctive question* represented by a formula $? \pm |A_1, \dots, A_n|$ that should be read: "Is it the case that A_1 and ... and is it he case that A_n?". To grasp the idea, consider binary conjunctive question: "Is it the case that A_1 and is it the case that A_2?". This question, abbreviated by $? \pm |A_1, A_2|$, has four direct answers: $A_1 \land A_2, A_1 \land \neg A_2, \neg A_1 \land A_2, \neg A_1 \land \neg A_2$. For precise definition of a conjunctive question see (Urbański, 2001, p. 76).

Example 5

$$?H(d)$$
$$H(d) \wedge F(w_1, d, Q_1) \rightarrow U(d, o_1, Q_1)$$
$$H(d) \wedge F(w_2, d, Q_2) \rightarrow U(d, o_2, Q_2)$$
$$\ldots$$
$$H(d) \wedge F(w_n, d, Q_n) \rightarrow U(d, o_n, Q_n)$$
$$F(w_1, d, Q_1) \wedge \ldots \wedge F(w_n, d, Q_n)$$
$$U(d, o_1, Q_1) \wedge \ldots \wedge U(d, o_n, Q_n) \rightarrow H(d))$$
$$? \pm |U(d, o_1, Q_1), \ldots, U(d, o_n, Q_n)|$$
$$?U(d, o_1, Q_1)$$

$$\begin{array}{cc} U(d, o_1, Q_1) & \neg U(d, o_1, Q_1) \\ ?U(d, o_2, Q_2) & \neg H(d) \end{array}$$

$$\begin{array}{cc} U(d, o_2, Q_2) & \neg U(d, o_2, Q_2) \\ \ldots & \neg H(d) \\ ?U(d, o_{n-1}, Q_{n-1}) & \end{array}$$

$$\begin{array}{cc} U(d, o_{n-1}, Q_{n-1}) & \neg U(d, o_{n-1}, Q_{n-1}) \\ ?U(d, o_n, Q_n) & \neg H(d) \end{array}$$

$$\begin{array}{cc} U(d, o_n, Q_n) & \neg U(d, o_n, Q_n) \\ H(d) & \neg H(d) \end{array}$$

To clarify the intuition of setting out the questions to be asked by some queries of an e-scenario, we will present some operationalisation. To do this, we assume that the interrogator will accept premises of the following form (here formulated in a first person manner):

(*) if d is a human and I formulate the condition w_i (as a task's condition) and then I ask d the question Q_i, then d gives back an answer o_i to the question Q_i.

In this scheme o_i is an answer to question Q_i such that (in the interrogator's opinion) exactly that kind of an answer would be given by a human being, taking condition w_i set for the task into account. We will write the scheme in symbols as the following:

(**) $H(d) \wedge F(w_i, d, Q_i) \rightarrow U(d, o_i, Q_i)$,

where $F(w_i, d, Q_i)$ stands for "I formulate the condition w_i for the task and then I ask d a question Q_i", and U, H are understood as before.

Let us assume that the interrogator uses n such premises (where $n > 1$). Then, the strategy for the interrogator is expressed by the e-scenario of example 5.

Due to this kind of approach, we obtain an easy way of differentiating the questions which the interrogator asks to himself/herself from the ones asked to the answerer. The first group of questions are: $?U(a, o_1, Q_1), \ldots, ?U(a, o_n, Q_n)$ (they are used in the e-scenario as implied questions), while the second one are: $Q_1, \ldots, Q_n$ (they are only mentioned in the e-scenario as the third arguments of the predicate U).

When we take the Golden Path Theorem into account, we may say that the judge carrying out the presented e-scenario will end the interrogation with accurate identification of the answerer. Of course, we should still have in mind that this would be possible if all declarative premises of the e-scenario were true (which is a rather strong assumption).

At this stage, one may clearly see that the final result of the TT relies heavily on the knowledge and beliefs of the interrogator (this is one of the serious issues of the TT's setting). An e-scenario ensures only that the interrogator will get the answer to the initial question of the e-scenario. This e-scenario, however, does not guarantee the true answer, which is understood as an accurate identification of the answerer. The identification's accuracy depends on the set of premises on the basis of which the interrogator builds his/her e-scenario for the TT. This consequence might be seen as a weak point of the TT setting. However, when we take a closer look, it appears that the problem has its roots much deeper, in the unclear and fuzzy criteria of "being a human" (cf. for example considerations on how we assign thinking to other human beings presented by Moor (1976) and Stalker (1976)).

It is worth noticing, that for the real-life Turing Test, at least some elements of reasoning involving probabilities would be necessary. Some additional statistical rules might be used e.g. to set a proportion of satisfactory to unsatisfactory answers obtained by **J**. We may however imagine that a procedure of defining these rules might be something like the proposal made by French (1996), i.e. the so-called *Human Subcognitive Profile*. According to French, it is possible to establish (using empirical procedures) the profile of human answers to questions concerning low level cognitive structures.

4 Conclusions and future work

Erotetic search scenarios allow for the formal modeling of an interrogator's hidden agenda. What is more they offer a possibility of a differentiation of questions posed by an interrogator to himself/herself from the ones that actually should be asked to an interlocutor. However, adequacy of resulting models is far from satisfactory. This is due to the fact that the semantics of underlying logic is rather simple and that assumed inferential relations (both declarative and erotetic ones) are relatively modest in character. Thus further development of this kind of formal models of interrogator's hidden agenda requires extension of the IEL framework to logics offering deeper insights into mechanisms underlying dialogue and argumentation.

References

Nuel D. Belnap. 1969. Åqvist corrections-accumulating question sequences. In: *Philosophical Logic*. J. W. Davis, P. J. Hockney, and K. Wilson (eds.), pp. 122–134. Reidel.

Paul Ekman. 2001. *Telling Lies: Clues to Deceit in the Marketplace, Politics, and Marriage*. W. W. Norton & Company, New York.

Robert French. 1996. The Inverted Turing Test: How a Mindless Program Could Pass It. *Psycholoquy*, 7(39).

Brooke Harrington (ed.). 2009. *Deception: From Ancient Empires to Internet Dating*. Stanford University Press.

Jan van Kuppevelt. 1995. Discourse structure, topicality and questioning. *Journal of Linguistics*, 31:109–147.

Paweł Łupkowski. 2010. *The Turing Test. Interrogator's Perspective* (*Test Turinga. Perspektywa sędziego*, in Polish). Adam Mickiewicz University Press, Poznań (in press).

James H. Moor. 1976. An Analysis of the Turing Test. *Philosophical Studies*, 30:249–257.

Raymond Smullyan. 1978. *What Is the Name of This Book*. Prentice-Hall, Englewood Cliffs, NJ.

Douglas F. Stalker. 1976. Why machines can't think: A reply to James Moor. *Philosophical Studies*, 34:317–320.

Mariusz Urbański. 2001. Synthetic tableaux and erotetic search scenarios: Extension and extraction. *Logique & Analyse*, 17–174–175:69–91.

Andrzej Wiśniewski. 1995. *The Posing of Questions: Logical Foundations of Erotetic Inferences*. Kluwer Academic Publishers, Dordrecht–Boston–London.

Andrzej Wiśniewski. 2003. Erotetic Search Scenarios. *Synthese*, 134(3):389–427.

Andrzej Wiśniewski. 2007. Constructive Erotetic Implication. (manuscript).

Andrzej Wiśniewski. 2010. Erotetic Logics. Internal Question Processing and Problem-solving. UNILOG'03, Lisbon, `http://www.staff.amu.edu.pl/~p_lup/aw_pliki/Unilog`.

Andrzej Wiśniewski and Jerzy Pogonowski. 2010. Interrogatives, Recursion, and Incompleteness. *Journal of Logic and Computation*, 10:114, in press (doi:10.1093/logcom/exq014).