Supporting Scientific Enquiry with Uncertain Sources

Federico Cerutti
Cardiff University, UK

Gavin Pearson
Dstl, UK

Abstract—In this paper we propose a computational methodology for assessing the impact of trust associated to sources of information in scientific enquiry activities building upon recent proposals of an ontology for situational understanding and results in computational argumentation. Often trust in the source of information serves as a proxy for evaluating the quality of the information itself, especially in the cases of information overhead. We show how our computational methodology, composed of an ontology for representing uncertain information and sources, as well as an argumentative process of conjecture and refutation, support human analysts in scientific enquiry, as well as highlighting issues that demand further investigation.

Index Terms—computational models of trust, scientific enquiry, uncertainty

I. INTRODUCTION

Individuals and organisations have access to a rich and diverse source of information that can be exploited for scientific enquiry—i.e. relating relevant information and form logical conclusions, as well as identifying gaps in information in order to answer a given query. However, an open and enduring problem resides in managing the impact of trust measurements in such activities.

We propose a computational methodology for assessing the impact of uncertainty in sources and information in scientific enquiry to support human analysts with a sound ontology and the ability to reason with partial information. In this paper we extend a previous proposal [5] to (1) not only to represent the trust in the sources, but also the quality of information provided by such sources; and (2) a computational model of handling inconsistencies and uncertain sources. We ground our preliminary investigation in a case study on the alleged links between vaccination and autism that will help us describing our desiderata for supporting human analysts (Section II). We then review our previous proposal [5] and other relevant literature in Section III. We propose a methodology satisfying those desiderata in Section IV and in Section V we present
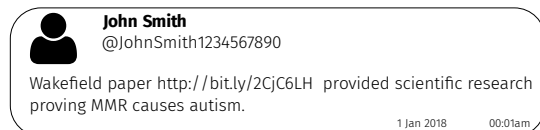
Fig. 1. A message on social media suggesting a research paper linking MMR vaccination and autism

contributions to support scientific enquiry with uncertain information and sources as part of the process of conjectures and refutation [17] embodied by formal argumentation with specific reference to our case study. As this preliminary work is part of an ongoing investigation, we will extensively discuss related work and future directions in Section VI.

II. MOTIVATIONAL SCENARIO AND DESIDERATA

Let us suppose that an individual—Alex—needs to answer the question "Do MMR vaccinations cause autism?" Like any other individual living at the beginning of the XXI century CE, Alex might rely first on the social networks of their choice. For instance, let us suppose that Alex encounters the message depicted in Figure 1,[1] where John shares a link to a research paper.

Alex follows the link and notices that it refers to a short conference abstract with little evidence attached. However, from a simple search on their favourite engine, Alex finds [24] (now retracted), where Wakefield et al. present an early report investigating the case of 12 children experiencing a loss of acquired skills, including language. The findings of the study are:

> Onset of behavioural symptoms was associated, by the parents, with measles, mumps, and rubella [MMR ed.] vaccination in eight of the 12 children, with measles infection in one child, and otitis media in another. [...] Behavioural disorders included autism (nine), disintegrative psychosis (one), and possible postviral or vaccinal encephalitis (two).

From this, Alex might infer—improperly applying a frequentist inference[2]—that [24] supports the conclusion that

[1] Although this message has been created by the authors, it is based on several messages that can be found on platforms like Twitter or Facebook. The authors chose not to display original social media messages as that would have been unethical.

[2] See [10] for instance on issues associated to statistical inferences and human psychology: further discussions on this topic are outside the scope of the present paper.

there is an association between MMR vaccination and autism with probability $\frac{8}{12} = 0.67$.

Alex also notices that [24] appears with a retraction notice [21] that states: "it has become clear that several elements of the 1998 paper by Wakefield et al are incorrect."

The Editors of the Lancet [21] suggest therefore that the methodology followed by [24] was not correct, without commenting on the findings. Hence, Alex searches for more recent papers that cite [24] and finds [15], whose findings are:

> Of the 537,303 children in the cohort (representing 2,129,864 person-years), 440,655 (82.0 percent) had received the MMR vaccine. We identified 316 children with a diagnosis of autistic disorder [...] There was no association between the age at the time of vaccination, the time since vaccination, or the date of vaccination and the development of autistic disorder.

Applying the same improper frequentist inference—on the data available in [15]—Alex now infers that [15] supports the conclusion that there is an association between MMR vaccination and autism with probability $\frac{316}{440,655} = 7.17\mathrm{e}-04$. Assuming the logical axiom of *tertium non datur* this leads Alex to imply that [15] supports the conclusion that there is *not* an association between MMR vaccination and autism with probability $\frac{440,655-316}{440,655} = 0.9992$.

In order to develop a computational machinery for supporting Alex's scientific enquiry with conflicting information and uncertain sources, we first need to identify principled desiderata. As identified in [5], (**Des1**) trust has to be context-dependent: the level of trust we have in @JohnSmith1234567890 when it comes to MMR vaccination should not impact the level of trust we have in @JohnSmith1234567890 on another topic, e.g. celebrities lifestyle.

Moreover, the second desideratum identified in [5] states that (**Des2**) trust needs to be expressed through an ordinal scale, i.e. it should be possible to determine whether, given the same context, a source of information is more trustworthy than another.

In this paper, we discuss two further desiderata. The third desideratum (**Des3**) states that information reported by sources should be linked to a measurement of uncertainty that needs to be expressed through an ordinal scale.

Finally, the fourth desideratum (**Des4**) is that information reported by sources should be subjected to a scientific enquiry process of conjectures and refutation. Also information related to the contextual trust in the source of information should not be exempt from such a process: it must be possible to criticise and refute the level of trust on a specific source.

III. Background

A. SitUTrustOnto

Figure 2, together with Table I, depicts SitUTrustOnto, the ontology of trust for situational understanding proposed in

TABLE I
ORL RULES IN SitUTrustOnto.

$$R1: \left\{ \begin{array}{l} \mathsf{Trust}(?x) \wedge \mathsf{DefaultTrust}(?y) \wedge \\ \mathsf{implementsDefault}(?x,?y) \wedge \\ \mathsf{TrustDescriptor}(?t) \wedge \\ \mathsf{hasDescriptor}(?y,?t) \end{array} \right\} \rightarrow \mathsf{hasDescriptor}(?x,?t)$$

$$R2: \left\{ \begin{array}{l} \mathsf{Source}(?x) \wedge \mathsf{Source}(?y) \wedge \\ \mathsf{contradicts}(?x,?y) \wedge \\ \mathsf{Trust}(?ax) \wedge \mathsf{Trust}(?ay) \wedge \\ \mathsf{hasTrustSameSource}(?x,?ax) \wedge \\ \mathsf{hasTrustSameSource}(?y,?ay) \\ \wedge \mathsf{Query}(?q) \wedge \mathsf{hasQuery}(?ax,?q) \wedge \\ \mathsf{hasQuery}(?ay,?q) \wedge \\ \mathsf{hasDescriptor}(?ay,?ty) \wedge \\ \mathsf{hasDescriptor}(?ax,?tx) \wedge \\ \mathsf{equalTrustThan}(?tx,?ty) \end{array} \right\} \rightarrow \mathsf{Issue}(?x)$$

$$R3: \left\{ \begin{array}{l} \mathsf{Source}(?x) \wedge \mathsf{Source}(?y) \wedge \\ \mathsf{contradicts}(?x,?y) \wedge \\ \mathsf{Trust}(?ax) \wedge \mathsf{Trust}(?ay) \wedge \\ \mathsf{hasTrustSameSource}(?x,?ax) \wedge \\ \mathsf{hasTrustSameSource}(?y,?ay) \\ \wedge \mathsf{Query}(?q) \wedge \mathsf{hasQuery}(?ax,?q) \wedge \\ \mathsf{hasQuery}(?ay,?q) \wedge \\ \mathsf{hasDescriptor}(?ay,?ty) \wedge \\ \mathsf{hasDescriptor}(?ax,?tx) \wedge \\ \mathsf{lowerTrustThan}(?tx,?ty) \end{array} \right\} \rightarrow \mathsf{Issue}(?y)$$

$R4: \{\mathsf{Issue}(?x) \wedge \mathsf{Source}(?y) \wedge \mathsf{contradicts}(?x,?y)\} \rightarrow \mathsf{Issue}(?y)$

[5].[3] An ontology can be described using tractable sub-sets of First Order Logic, which take the name of Description Logics. In the following we will be using a specific Description Logic named $\mathcal{ALI}+$.

An ontology comprises two components [1]: the vocabulary (TBox) and the assertions about individuals (ABox). The vocabulary consists of concepts, which denote sets of individuals; and roles, which denote binary relationships between individuals.

Elementary descriptions are atomic concepts and atomic roles. Complex descriptions can be built from them inductively with concept constructors. In abstract notation, we use the letters A and B for atomic concepts, the letter R for atomic roles, and the letters C and D for concept descriptions. Concept descriptions in $\mathcal{ALI}+$ are formed according to the following syntax rule: $\mathsf{C}, \mathsf{D} \rightarrow \mathsf{A}$ (atomic concept) $\mid \top$ (universal concept) $\mid \perp$ (bottom concept) $\mid \mathsf{C} \sqcap \mathsf{D}$ (intersection) $\mid \forall R.\mathsf{C}$ (value restriction, or range) $\mid \exists R.\top$ (limited existential quantification, or domain). R^- denotes the inverse role; roles can be transitive and symmetric. The syntax of $\mathcal{ALI}+$ can clearly be presented using the OWL 2 Web Ontology Language.[4] Semantics are given—as usual—by means of an interpretation.

For ease of representation, in the following we will use a graph-based syntax, where nodes are either atomic concepts or individuals (identified by ♦), and edges are labelled with

---

[3] Available at https://goo.gl/DLq6dz
[4] https://www.w3.org/TR/owl2-overview/

(transitive) higherTrustThan ≡
(transitive) lowerTrustThan⁻

equalTrustThan

Issue

(symmetric) equalTrustThan   (symmetric) incomparable

TrustDescriptor ← type — ♦ CompletelyReliable

is-a

higherTrustThan
equalTrustThan

type   type

♦ UsuallyReliable

Source   contradicts

higherTrustThan
equalTrustThan

hasSource ≡
hasTrustSameSource⁻

type

♦ FairlyReliable

Query

hasDescriptor ≡
hasTrustSameDescriptor⁻

higherTrustThan
equalTrustThan

hasQuery ≡
hasTrustSameQuery⁻

♦ NotUsuallyReliable

incomparable

Trust

higherTrustThan
equalTrustThan

incomparable   equalTrustThan

is-a   implementsDefault

♦ Unreliable ← incomparable — ♦ CannotBeJudged

DefaultTrust
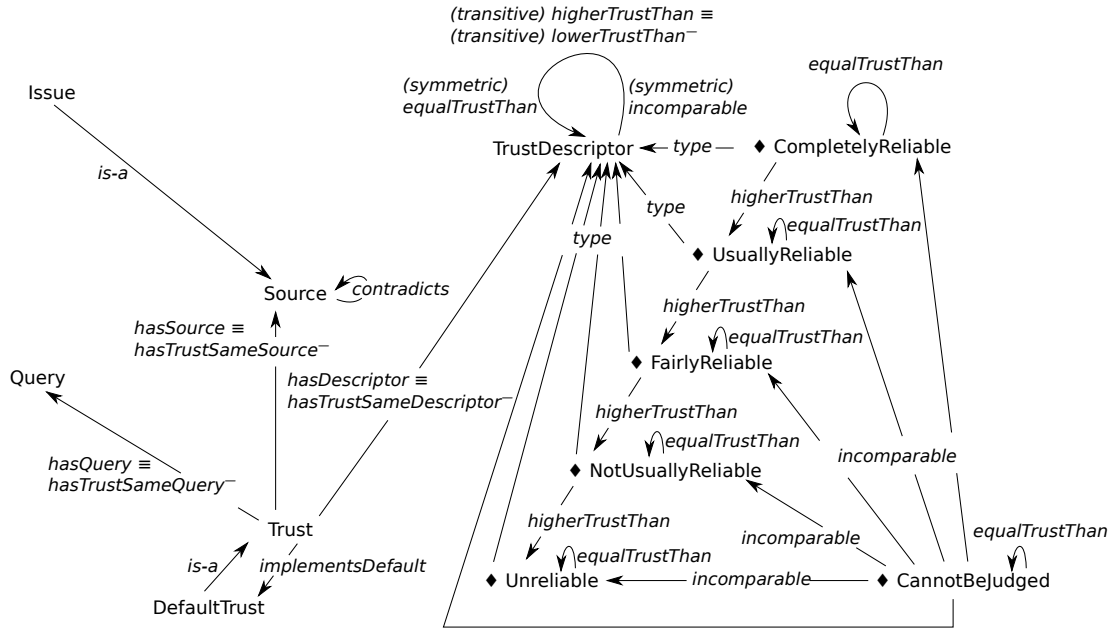
Fig. 2. Graphical representation of SitUTrustOnto: nodes with ♦ are individuals, otherwise atomic concept. Edges are labelled with *is-a* (subclass relations), with *type* (membership assertion); and with atomic roles. For instance, *(transitive) higherTrust ≡ (transitive)lowerTrustThan⁻* describes the role *higherTrust* with domain and range Trust, such as it is the inverse of *lowerTrust*. Issue represents the class of Sources that do require further investigation as they are deeemed to be problematic.

Reliability of the Source		Credibility of the Information	
A	Completely reliable	1	Confirmed by other sources
B	Usually reliable	2	Probably true
C	Fairly reliable	3	Possibly true
D	Not usually reliable	4	Doubtful
E	Unreliable	5	Improbable
F	Reliability cannot be judged	6	Truth cannot be judged

one of the following labels: *is-a* (representing sub-concept relation), *type* (membership assertion), or *R* (roles).

Moreover, as presented in [12], ORL is a language for expressing Horn clause rules that extends the OWL language in a syntactically and semantically coherent manner. A rule has the form *antecedent → consequent*, where both *antecedent* and *consequent* are conjunctions of atoms written $a_1 \wedge \ldots \wedge a_n$. Variables are prefixed with a question mark—e.g. ?x. The model-theoretic semantics for ORL is an extension of the semantics given by an interpretation. A rule is satisfied by an interpretation iff every binding—mapping to elements of the domain—that satisfies the antecedent also satisfies the consequent. An interpretation satisfies an ontology iff it satisfies every axiom (including rules) and fact in the ontology [12].

The concept Source in Figure 2 contains sources of information, e.g. blog posts, twits, scientific papers.... The concept Query describes the situation that needs to be understood, e.g. specific queries an analyst has to address, such as "Do vaccinations cause autism?"

The concept TrustDescriptor contains trust descriptors: given

our interest in supporting human analysts and not to replace them, we chose to rely on the *admiralty code* [18] that identifies the six descriptors listed in Table II (left). Figure 2 shows the six individuals belonging to TrustDescriptor, and their relationships expressed by *equalTrustThan*—identifying that two descriptors are equivalent; *higherTrustThan* and *lowerTrustThan*—expressing ordering, and thus satisfying (**Des2**); and *incomparable*—when two descriptors cannot be compared.

The concept Trust describes the relationship between a source of information, a query, and a trust descriptor, thus providing the context in which to assess the trust in a source of information for a given query. Please note that there is no role linking directly Source to TrustDescriptor, nor Source to Query, thus satisfying (**Des1**). Moreover, DefaultTrust is a sub-concept of Trust that provides *default* trust accounts between some types of queries and some sources of information. For instance, we might want to express that scientific papers addressing medical queries generally have high trust, and thus we can create a specific individual belonging to DefaultTrust. This means that when a new piece of information is added to the ontology, we can exploit defaults for assessing its trustworthiness using the rule R1 from Table I.

Finally, since different Sources can contradict each other, Issue is populated by the means of rules R2, R3, and R4 of Table I. This is the case where two sources of information ♦ a and ♦ b, related to the same query, contradict each other, i.e. ♦ a *contradicts* ♦ b, and either they have the same level of trust (*equalTrustThan*), cf. R2 in Table I; or the trust in ♦ a is *lowerTrustThan* ♦ b, cf. R3 in Table I. This is based

on the assumption that if ♦ a *contradicts* ♦ b, but ♦ a is more trustworthy than ♦ b, an analyst would accept ♦ a and discard ♦ b. Finally, if a source of information also belong to Issue, also the sources it is in conflict with belong to Issue (cf. R4 in Table I.

B. Handling conflicting information through argumentation

A widely used approach to handle uncertain and conflicting information is argumentation theory, where arguments can represent reasons and inferences in favours of conclusions. Among other approaches, in deductive argumentation each argument is defined using a logic, and in the following we adopt the simple, but elegant logic proposed in [2]. Thus, we let $\mathcal{L}$ be a logical language. If α is an atom in $\mathcal{L}$, then α is a *positive literal* in $\mathcal{L}$, and $\neg\alpha$ is a *negative literal* in $\mathcal{L}$. For a literal β, the *complement* of the positive literal $\beta = \alpha$ is $\overline{\beta} = \neg\alpha$ (resp. if $\beta = \neg\alpha$ is not a positive literal, its complement is the positive literal $\overline{\beta} = \alpha$).

A *simple rule* is of the form $\alpha_1 \wedge \ldots \wedge \alpha_k \to \beta$ where $\alpha_1, \ldots, \alpha_k, \beta$ are literals. A *simple logical knowledge base* is a set of literals and a set of simple rules. Given a simple logic knowledge base, Δ, the *simple consequence relation* $\vdash_s$ is defined, such that $\Delta \vdash_s \beta$ if and only if there is a rule $\alpha_1 \wedge \ldots \wedge \alpha_n \to \beta \in \Delta$ and $\forall i$ either $\alpha_i \in \Delta$ or $\Delta \vdash_s \alpha_i$. Now, given $\Phi \subseteq \Delta$ and a literal α, $\langle \Delta, \alpha \rangle$ is a *simple argument* if and only if $\Phi \vdash_s \alpha$ and $\nexists \Phi' \subsetneq \Phi$ such that $\Phi' \vdash_s \alpha$. Φ is the support (or premises, assumptions) of the argument, and α is the claim (or conclusion) of the argument. Given an argument $\mathbf{a} = \langle \Phi, \alpha \rangle$, the function Support($\mathbf{a}$) returns Φ, and Claim($\mathbf{a}$) returns α.

For simple arguments $\mathbf{a}$ and $\mathbf{b}$ we consider the following types of simple attack:

- $\mathbf{a}$ is a *simple undercut* of $\mathbf{b}$ if there is a simple rule $\alpha_1 \wedge \alpha_k \to \beta$ in Support($\mathbf{b}$) and there is an $\alpha_i \in \{\alpha_1, \ldots, \alpha_k\}$ such that Claim($\mathbf{a}$) is the complement of α_i;
- $\mathbf{a}$ is a *simple rebut* of $\mathbf{b}$ if Claim($\mathbf{a}$) is the complement of Claim($\mathbf{b}$).

Some arguments can be instances of argumentation schemes [25], [27], viz. abstract reasoning patterns commonly used in everyday conversational argumentation, legal, scientific argumentation, etc. Schemes have been derived from empirical studies of human argument and debate. They can capture traditional deductive and inductive approaches as well as plausible reasoning. Each scheme has a set of critical questions that represents standard ways of critically probing into an argument to find aspects of it that are open to criticism.

For instance, the following is the argumentation scheme for argument from correlation to cause:

Correlation Premise: There is a positive correlation between A and B.

Conclusion: A causes B.

Critical questions are:

CQ1: Is there really a correlation between A and B?
CQ2: Is there any reason to think that the correlation is any more than a coincidence?

CQ3: Could there be some third factor, C, that is causing both A and B?

For the purpose of this paper, let us also recall the argumentation scheme for argument from witness testimony:

Position to Known Premise: Witness W is in a position to know whether A is true or not
Truth Telling Premise: Witness W is telling the truth (as W knows it).
Statement Premise: Witness W states that A is true (false).
Conclusion: Therefore, A may be plausibly taken to be true (false).

with associated the following critical questions:

CQ1: Is what the witness said internally consistent?
CQ2: Is what the witness said consistent with the known facts of the case (based on evidence apart from what the witness testified to)?
CQ3: Is what the witness said consistent with what other witnesses have (independently) testified to?
CQ4: Is there some kind of bias that can be attributed to the account given by the witness?
CQ5: How plausible is the statement A asserted by the witness?

The Argument Interchange Format+ (AIF+) [7], [19], [20][5] is the current proposal for a standard notation for argument structures. It is based on a graph that specifies two types of nodes: information nodes (or I-nodes) and scheme nodes (or S-nodes). These are represented by two disjoint sets, $\mathcal{N}_I \cup \mathcal{N}_S = \mathcal{N}$ and $\mathcal{N}_I \cap \mathcal{N}_S = \varnothing$, where information nodes represent claims, premises, data, etc., and scheme nodes capture the application of patterns of reasoning belonging to a set $\mathcal{S} = \mathcal{S}^R \cup \mathcal{S}^C \cup \mathcal{S}^P \cup \mathcal{S}^Y$ with $\mathcal{S}^R$, $\mathcal{S}^C$, $\mathcal{S}^P$, $\mathcal{S}^Y$ pairwise disjoint. Reasoning patterns can be of three types: rule of inference $\mathcal{S}^R$; criteria of preference $\mathcal{S}^P$; criteria of conflicts $\mathcal{S}^C$; and criteria of illocutions $\mathcal{S}^Y$.

The relation fulfils $\subseteq \mathcal{N}_S \times \mathcal{S}$ expresses that a scheme node instantiates a particular scheme. Scheme nodes, moreover, can be one of four types: rule of inference application nodes $\mathcal{N}_S^{RA}$; preference application nodes $\mathcal{N}_S^{PA}$; conflict application nodes $\mathcal{N}_S^{CA}$; or illocutions nodes $\mathcal{N}_S^{YA}$, with $\mathcal{S} = \mathcal{N}_S^{RA} \cup \mathcal{N}_S^{PA} \cup \mathcal{N}_S^{CA} \cup \mathcal{N}_S^{YA}$, and $\mathcal{N}_S^{RA}$, $\mathcal{N}_S^{PA}$, $\mathcal{N}_S^{CA}$, $\mathcal{N}_S^{YA}$ pairwise disjoint.

From a graphical perspective, rectangular nodes represent information nodes, while circle ones represent scheme nodes: black text on white for RA nodes, and white text on black for CA nodes.

Computing the acceptability of arguments is the main scope of using an argumentation framework [9] that consists of a set of arguments and a binary attack relation between them.

Definition 1: An *argumentation framework* (AF) is a pair $\Gamma = \langle \mathcal{A}, \mathcal{R} \rangle$ where $\mathcal{A}$ is a set of arguments and $\mathcal{R} \subseteq \mathcal{A} \times \mathcal{A}$. We say that $\mathbf{b}$ *attacks* $\mathbf{a}$ iff $\langle \mathbf{b}, \mathbf{a} \rangle \in \mathcal{R}$, also denoted as $\mathbf{b} \to \mathbf{a}$.

---

[5]In the following we will consider a formalisation based upon the OWL AFI+ ontology available at http://www.arg.dundee.ac.uk/wp-content/uploads/AIF.owl (on 8 Mar 2018).

Each argumentation framework, therefore, has an associated directed graph where the vertices are the arguments, and the edges are the attacks.

The basic properties of conflict–freeness, acceptability, and admissibility of a set of arguments are fundamental for the definition of argumentation semantics.

Definition 2: Given an *AF* $\Gamma = \langle \mathcal{A}, \mathcal{R} \rangle$:

- a set $S \subseteq \mathcal{A}$ is a *conflict–free* set of Γ if $\nexists$ $\mathbf{a}, \mathbf{b} \in S$ s.t. $\mathbf{a} \to \mathbf{b}$;
- an argument $\mathbf{a} \in \mathcal{A}$ is *acceptable* with respect to a set $S \subseteq \mathcal{A}$ of Γ if $\forall \mathbf{b} \in \mathcal{A}$ s.t. $\mathbf{b} \to \mathbf{a}$, $\exists$ $\mathbf{c} \in S$ s.t. $\mathbf{c} \to \mathbf{b}$;
- a set $S \subseteq \mathcal{A}$ is an *admissible* set of Γ if S is a conflict–free set of Γ and every element of S is acceptable with respect to S, i.e. $S \subseteq \mathcal{F}_\Gamma(S)$.

An argumentation semantics σ prescribes for any *AF* Γ a set of *extensions*, denoted as $\mathcal{E}_\sigma(\Gamma)$, namely a set of sets of arguments satisfying the conditions dictated by σ. For instance, here is the definition of preferred (denoted as PR) semantics.

Definition 3: Given an *AF* $\Gamma = \langle \mathcal{A}, \mathcal{R} \rangle$, a set $S \subseteq \mathcal{A}$ is a *preferred extension* of Γ, i.e. $S \in \mathcal{E}_{PR}(\Gamma)$, iff S is a maximal (w.r.t. set inclusion) admissible set of Γ.

Given a semantics σ, an argument $\mathbf{a}$ is said to be *credulously accepted* w.r.t. σ if $\mathbf{a}$ belongs to at least one σ-extension. $\mathbf{a}$ is *skeptically accepted* w.r.t. σ if $\mathbf{a}$ belongs to all the σ-extensions.

IV. SitUTrustOnto+: DEALING WITH UNCERTAINTY OF INFORMATION

Figure 3 depicts SitUTrustOnto+, a complete ontology for dealing with uncertainty of information. It extends SitUTrustOnto as it becomes manifest by a comparison with Figure 2. It allows for uncertainty in information to be expressed over an ordinal scale, implementing the right side of Table II, cf. the UncertaintyDescriptor in the figure. This satisfies (**Des3**).

Moreover, ensuring that both Illocution and Trust are subclasses of Node allows to express uncertainty about the assessment of the degree of contextual trust of a source of information as well as about the illocutionary statements between sources and information, thus satisfying (**Des4**). It is worth noticing that because Illocution is a subclass of Node, illocutionary acts can refer to other illocutionary acts with different degrees of uncertainty. In order to facilitate future integration with AIF+, Scheme Node $\equiv \mathcal{N}_S$; Information Node $\equiv \mathcal{N}_I$; and Illocution $\equiv \mathcal{N}_S^{YA}$.

Let us now illustrate how SitUTrustOnto+ enables us to represent the case study we presented in Section II.

First of all, let us summarise with ♦ MMR_Causes_Autism (of type Information Node) an individual representing the sentence depicted in Figure 1, whose language seems to suggest that *hasDescriptor*(♦ MMR_Causes_Autism, ♦ Probably true). This relates to ♦ JS_MMR (type Illocution) in a way that *IllocutionaryContent*(♦ JS_MMR, ♦ MMR_Causes_Autism) and *hasSource*(♦ JS_MMR, ♦ @JohnSmith1234567890).

Therefore the Illocution ♦ JS_MMR informs us that the Source ♦ @JohnSmith1234567890 asserts the *IllocutionaryContent* ♦ MMR_Causes_Autism. Assuming that @JohnSmith1234567890 does not reveal any affiliation to university of research centre, it appears that ♦ JS_autism_trust (type Trust) is such that *hasSource*(♦ JS_autism_trust, ♦ @JohnSmith1234567890); *hasQuery*(♦ JS_autism_trust, ♦ MedicalDomainVaccinations); *hasDescriptor*(♦ JS_autism_trust, ♦ CannotBeJudged).

Let us consider the main finding of [24], that represents a case of an illocutionary act referring to another illocutionary act. Indeed, the main piece of information is ♦ MMR_Associated_Autism ("Onset of behavioural symptoms was associated [...] with measles, mumps, and rubella [MMR ed.] vaccination") of type Information Node. This piece of information, according our simplified analysis in Section II, has probability 0.67 that, according to [18], maps into a *hasDescriptor*(♦ MMR_Associated_Autism, ♦ Possibly true). Such a piece of information is contained in an illocutionary act ♦ P_MMR such that *IllocutionaryContent*(♦ P_MMR, ♦ MMR_Associated_Autism) and *hasSource*(♦ P_MMR, ♦ Parents), with an associated trust assessment ♦ P_autism_trust where *hasSource*(♦ P_autism_trust, ♦ Parents); *hasQuery*(♦ P_autism_trust, ♦ MedicalDomainVaccination); *hasDescriptor*(♦ P_autism_trust, ♦ FairlyReliable) (although parents might not have a medical degree, they might best positioned to identify first symptoms).

We can now treat [24] as a source *per se*, hence having ♦ W_P_MMR (type Illocution) such that *IllocutionaryContent*(♦ W_P_MMR, ♦ P_MMR), and *hasSource*(♦ W_P_MMR, ♦ Wakefield et al., 1998). Moreover, its trust assessment ♦ W_P_MMR_trust— given that [24] has been published on The Lancet, a prestigious medical journal—should be such that *hasSource*(♦ W_P_MMR_trust, ♦ Wakefield et al., 1998); *hasQuery*(♦ W_P_MMR_trust, ♦ MedicalDomainVaccination); *hasDescriptor*(♦ W_P_MMR_trust, ♦ CompletelyReliable).

Finally, let us consider [15] and its main finding ♦ No_Associations (type Information node) that comes with *hasDescriptor*(♦ No_Associations, ♦ Confirmed by other sources) according to the naïve probabilistic evaluation in Section II and [18], after the approximation of 0.9994 to 1.0 (1:1666 error). In addition, there are two symmetric instances of the conflicting scheme Logical Conflict between ♦ No_Associations and ♦ MMR_Associated_Autism, cf. [7]. ♦ No_Associations comes to the illocutionary act ♦ M_MMR (type Illocution) such that *IllocutionaryContent*(♦ M_MMR, ♦ No_Associations) and *hasSource*(♦ M_MMR, ♦ Madsen et al., 2002). Similarly as per [24], [15] has been published in a respectable medical journal, hence its trust assessment ♦ M_MMR_trust is such that *hasSource*(♦ M_MMR_trust, ♦ Madsen et al., 2002); *hasQuery*(♦ M_MMR_trust, ♦ MedicalDomainVaccination); *hasDescriptor*(♦ M_MMR_trust, ♦ CompletelyReliable).
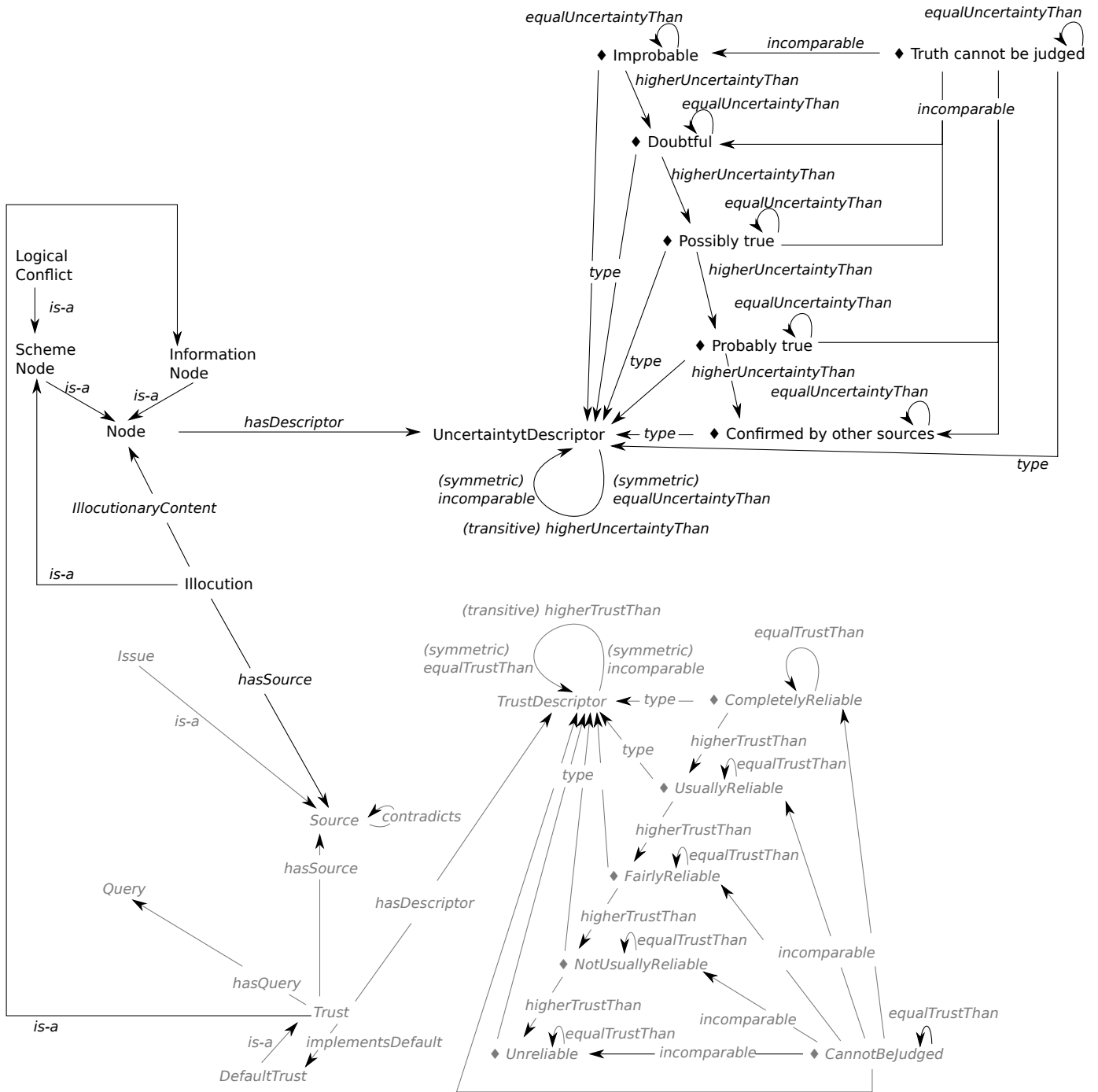
Fig. 3. Graphical representation of SitUTrustOnto+: nodes with ♦ are individuals, otherwise atomic concept. Edges are labelled with *is-a* (subclass relations), with *type* (membership assertion); and with atomic roles. Differently from Figure 2, inverse relationships have been omitted for improving the readability. The parts in common with Figure 2 are depicted in grey.
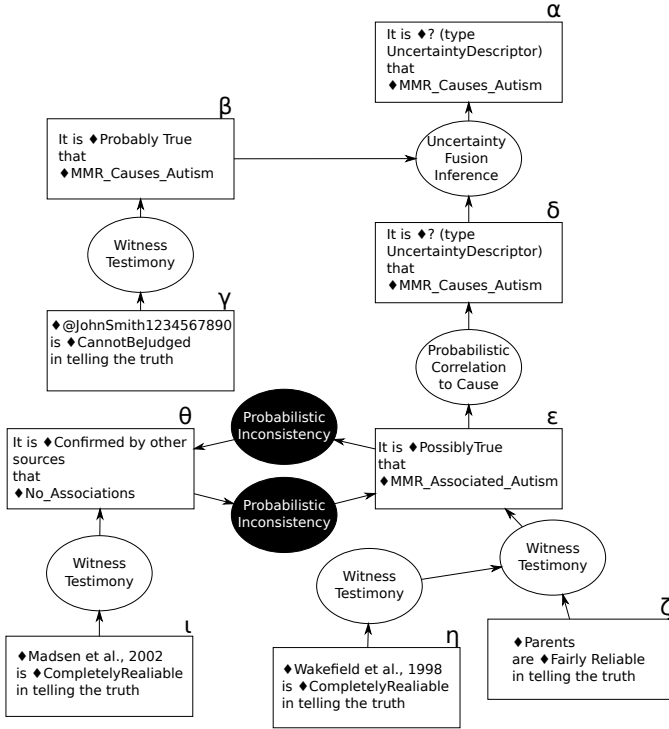
Fig. 4. Networks depicting the formal assessment of arguments considered in the case study discussed in Section II. Greek letters are shortcut labels to ease references in the text.

V. MANAGING CONFLICTS IN INFORMATION AND TRUST ON SOURCES

Figure 4 provides a snapshot of the process of conjectures and refutation [17] at the basis of scientific enquiry, with formal account of uncertainty on claimed pieces of information as well as variable degrees of trustworthiness associated to sources. It also illustrates the main contribution of our study—in addition to the SitUTrustOnto+ ontology described in Section IV—namely: (1) a conceptual shift towards having arguments about uncertainty assessment of information, which also require to consider probabilistic inconsistencies as conflict schemes; (2) methodologies for generating arguments with fused uncertainty assessments of information explicitly represented as argumentative inferences; (3) representing illocutionary acts as arguments instances of a modified, recursive, version of the argumentation scheme from witness testimony.

As per the conceptual shift towards having arguments about uncertain, let us consider statements θ and ϵ. θ states that "It is ♦ Confirmed by other sources that ♦ No_Associations" while ϵ that "It is ♦ PossiblyTrue that ♦ MMR_Associated_Autism," and from Section IV we also know that ♦ No_Associations = ¬♦ MMR_Associated_Autism. Hence this is a case where there is a probabilistic inconsistency between these two statement [4] as both their uncertainty assessment are quite high (greater than 0.5) and they are logically conflicting. This approach is rather different from current views in probabilistic argumentation [13], where there are two main families of approaches to probabilistic argumentation: constellation ap-

proaches, where arguments and attacks are uncertain; and epistemic approaches, where there is uncertainty about whether an argument is believed. We will discuss connections with that literature in future work.

As per representing fusion of uncertain information, let us consider the statements α, β, and δ. They all refer to ♦ MMR_Causes_Autism, but they serve three distinct epistemic purposes. β represents the illocutionary act performed by ♦ @JohnSmith1234567890; δ represents the results of a statistical inference based upon a (undefined, but rather straightforward to derive) argumentation scheme from probabilistic correlation to cause; and α is the result of a fusion operator over β and δ, in the case a final summary uncertainty assessment is necessary, and under the assumption that such as fusion depends only on β and δ: more articulated fusion operators perhaps taking into consideration the totality of evidence can be envisaged, and we will explore them in future work.

Regarding representing illocutionary acts as arguments themselves, we rely on an evolution of the argumentation scheme from witness testimony, where the truth telling premise is modified into "Witness W is telling the truth with reliability assessment T:" γ, ζ, η, ι are examples of such statements. Such an evolution is inspired by [16], where trust is explicitly mentioned in argumentation schemes, albeit not represented with measures of uncertainty.

In addition, the discussion in Section II suggests the need for an extended AIF+ model, where sources of information can witness illocutionary acts of other sources, as it is the case of ♦ Wakefield et al., 1998 witnessing the illocutionary act of ♦ Parents. From a sintactic point of view, Figure 4 is not compliant with AIF+: it also raises questions on the semantics of the logical relationships we are representing there. Indeed if η is the premise of an instance of the argumentation scheme from witness testimony, it is immediate to conclude that there must be a conclusion for such an argument, hence we can introduce the equivalent of a *Skolem term* [14]. For instance, using simple logic (cf. Section III), we would introduce a new term τ and write the following two rules $(\eta \rightarrow \tau)$, and $(\zeta \wedge \tau \rightarrow \epsilon)$. We will explore further these aspects in future work.

Linked to the above point, also the acceptability status of argument needs to be reconsidered. Indeed, from Figure 4, we can derive the following two arguments: $\mathbf{a} = \langle \{\iota, \iota \rightarrow \delta\}, \delta \rangle$, and $\mathbf{b} = \langle \{\eta, \zeta, \eta \rightarrow \tau, \zeta, \tau \rightarrow \epsilon\}, \epsilon \rangle$ with $\delta = \bar{\epsilon}$ and $\epsilon = \bar{\delta}$, hence having $\mathbf{a}$ and $\mathbf{b}$ mutually conflicting (simple rebut). Questions now relate to (1) how to derive a meaningful measure of uncertainty for the arguments as a whole (e.g. using discounting operators); and (2) the effects on trustworthiness of sources as well as uncertainty assessments considered in arguments deemed to be unacceptable according to an argumentation semantics (cf. Section III-B), hence providing specific semantics to the (**Des4**), sintactically addressed in Section IV. It is also worth mentioning that a meaningful function deriving the strength of an argument should not be based only upon the uncertainty assessment of information and of sources of information, but also on the quality of the argumentation itself including answers to critical questions—

i.e. an argument should not be considered warranted unless reasonable answers to each critical question are provided [27]—and reasonable standards for burden of proof [26].

VI. Conclusion

In this paper we proposed a computational methodology for evaluating trust associated to sources of information in scientific enquiry. In particular, we expand on [5] to present a complete ontology for dealing with uncertainty in information and in sources of information thus satisfiying the four desiderata discussed in Section II. Moreover, since in scientific enquiry—as well as in other settings such as situational understanding, i.e. CISpaces.org [22]—pieces of information links together to form arguments, we explicitly link the AIF+ [7], [19], [20] ontology. We also present novel extensions to formal argumentation theories, in particular: (1) a conceptual shift towards having arguments about uncertainty assessment of information, which also require to consider probabilistic inconsistencies as conflict schemes; (2) methodologies for generating arguments with fused uncertainty assessments of information explicitly represented as argumentative inferences; and (3) representing illocutionary acts as arguments themselves, in particular as instances of a modified, recursive version of the argumentation scheme from witness testimony [27].

As part of future work we plan to evaluate techniques for automatic evaluation of trust: we will investigate how topic modelling—i.e. automatically identifying relevant topics in a written document, for instance using the Latent Dirichlet Allocation (LDA) [3]—and similarities of sources—e.g. articles in The Lancet are more similar to articles in the New England Journal of Medicine rather than to articles in the Daily Mail—can help suggesting trust measures for new pieces of information. Moreover, we will also investigate how to automatically identify problems with default assumptions, i.e. questioning whether there is enough evidence suggesting that a default assignment needs to be revisited.

The ultimate goal of our research is to support any analyst in their scientific enquiry. We will therefore consider natural language interfaces to formal argumentation [6], also building on top of the CISpaces.org project, as well as conversational interfaces based on controlled natural language, e.g. [23], to better support analysts in their activities.

In future work we will compare with other approaches to argue about trustworthiness of sources as recently presented in literature, such as [11]. Moreover, we will also consider similarities and limitations of this approach w.r.t. approaches for qualitative decision under uncertainty [8].

References

[1] F. Baader and W. Nutt. Basic Description Logics. In *The Description Logic Handbook*, pages 43–95. Cambridge University Press, 2003.

[2] P. Besnard and A. Hunter. Constructing argument graphs with deductive arguments: a tutorial. *Argument & Computation*, 5(1):5–30, 2014.

[3] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3:993–1022, 2003.

[4] G. D. Bona, M. Finger, N. Potyka, and M. Thimm. Inconsistency measurement in probabilistic logic. In J. Grant and M. V. Martinez, editors, *Measuring Inconsistency in Information*, volume 73 of *Studies in Logic*. College Publications, February 2018.

[5] O. Carpanini and F. Cerutti. Towards an Ontology of Trust for Situational Understanding. In *Advances in Intelligent Systems and Computing*, volume 650, 2018.

[6] F. Cerutti, N. Tintarev, and N. Oren. Formal Arguments, Preferences, and Natural Language Interfaces to Humans: an Empirical Evaluation. In *21st European Conference on Artificial Intelligence*, pages 207–212, 2014.

[7] C. I. Chesnevar, J. McGinnis, S. Modgil, I. Rahwan, C. Reed, G. R. Simari, M. South, G. A. W. Vreeswijk, and S. Willmot. Towards an argument interchange format. *The Knowledge Engineering Review*, 21(04):293, 2006.

[8] D. Dubois, H. Prade, and A. Rico. Residuated variants of Sugeno integrals: Towards new weighting schemes for qualitative aggregation methods. *Information Sciences*, 329:765–781, 2016.

[9] P. M. Dung. On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming, and n-Person Games. *Artificial Intelligence*, 77(2):321–357, 1995.

[10] G. Gigerenzer. The Psychology of Good Judgment. *Medical Decision Making*, 16(3):273–280, aug 1996.

[11] S. Gottifredi, L. H. Tamargo, A. J. Garca, and G. R. Simari. Arguing about informant credibility in open multi-agent systems. *Artificial Intelligence*, pages –, 2018.

[12] I. Horrocks and P. F. Patel-Schneider. A proposal for an OWL rules language. In *Proceedings of WWW '04*, pages 723–731, 2004.

[13] A. Hunter. A probabilistic approach to modelling uncertain logical arguments. *International Journal of Approximate Reasoning*, 54(1):47–81, 2013.

[14] S. C. Kleene. *Introduction to metamathematics*. Wolters-Noordhoff Pub, 6th edition, 1971.

[15] K. M. Madsen, A. Hviid, M. Vestergaard, D. Schendel, J. Wohlfahrt, P. Thorsen, J. Olsen, and M. Melbye. A Population-Based Study of Measles, Mumps, and Rubella Vaccination and Autism. *New England Journal of Medicine*, 347(19):1477–1482, nov 2002.

[16] S. Parsons, K. Atkinson, Z. Li, P. McBurney, E. Sklar, M. Singh, K. Haigh, K. Levitt, and J. Rowe. Argument schemes for reasoning about trust. *Argument & Computation*, 5(2-3):160–190, 2014.

[17] K. R. Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Routledge, 5th edition, 1989.

[18] H. Prunckun. *Handbook of Scientific Methods of Inquiry for Intelligence Analysis*. The Scarecrow Press, 2010.

[19] I. Rahwan and C. Reed. The Argument Interchange Format. In *Argumentation in Artificial Intelligence*, pages 383–402. Springer US, Boston, MA, 2009.

[20] C. Reed, S. Wells, J. Devereux, and G. Rowe. Aif+: Dialogue in the argument interchange format. In *Proceedings of the 2008 Conference on Computational Models of Argument: Proceedings of COMMA 2008*, pages 311–323, Amsterdam, The Netherlands, The Netherlands, 2008. IOS Press.

[21] The Editors of The Lancet. Retraction–Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, 375(9713):445, 2010.

[22] A. Toniolo, T. J. Norman, A. Etuk, F. Cerutti, R. W. Ouyang, M. Srivastava, N. Oren, T. Dropps, J. A. Allen, and P. Sullivan. Agent Support to Reasoning with Different Types of Evidence in Intelligence Analysis. In *Proceedings of AAMAS 2015*, pages 781–789, 2015.

[23] A. Toniolo, A. D. Preece, W. Webberley, T. J. Norman, P. Sullivan, and T. Dropps. Conversational intelligence analysis. In *Proceedings of the 17th International Conference on Distributed Computing and Networking*, ICDCN '16, pages 42:1–42:6, New York, NY, USA, 2016. ACM.

[24] A. Wakefield, S. Murch, A. Anthony, J. Linnell, D. Casson, M. Malik, M. Berelowitz, A. Dhillon, M. Thomson, P. Harvey, A. Valentine, S. Davies, and J. Walker-Smith. Ileal-lymphoid-nodular hyperplasia, non-specific colitis, and pervasive developmental disorder in children. *The Lancet*, 351(9103):637–641, 1998.

[25] D. Walton. Argumentation Theory: A Very Short Introduction. In *Argumentation in Artificial Intelligence*, pages 1–22. Springer US, Boston, MA, 2009.

[26] D. Walton. *Burden of Proof, Presumption and Argumentation*. Cambridge University Press, New York, NY, USA, 2014.

[27] D. Walton, C. Reed, and F. Macagno. *Argumentation schemes*. Cambridge University Press, NY, 2008.