Two Party Immediate Response Disputes: Properties and Efficiency [*]

Paul E.Dunne and T.J.M.Bench-Capon
Dept. of Computer Science, University of Liverpool,Liverpool L69 7ZF, United Kingdom

Abstract. Two Party Immediate Response Disputes (TPI-disputes) are one class of *dialogue* or *argument game* in which the protagonists take turns producing counter arguments to the 'most recent' argument advanced by their opponent. Argument games have been found useful as a means of modelling dialectical discourse and in providing semantic bases for proof theoretic aspects of reasoning. In this article we consider a formalisation of TPI-disputes in the context of finite *Argument Systems*. Our principal concern may, informally, be phrased as follows: given a specific argument system, $\mathcal{H}$ and argument, x within $\mathcal{H}$, what can be stated concerning the number of rounds a dispute might take for one of its protagonists to accept that x has *some* defence respectively *cannot* be defended?

Keywords: Argument Systems, Dialogue Game, Gentzen System, Proof Complexity

1. Introduction

In this paper we are concerned with two important formalisms that have been the subject of much interest with respect to their application in modelling dialectical process: *Argument Systems* [15], and *Argument Games* [19, 26]. Our principal concern is with the *length* of disputes when they are conducted in accordance with the etiquette prescribed by a particular formal protocol. The protocol of interest - TPI-dispute - was outlined in the work of [34] and in Section 1.2 we present a rigorous formalisation of this with examples of its operation being described in Section 2. The main technical concerns are dealt with in Section 3, wherein two questions are examined. Informally, these may be viewed as follows: suppose we are presented with an argument system and an argument within this. If it is required to observe the dispute rules prescribed in some dispute protocol,

a) when the given argument *can* be defended, how many rounds *could* it take to prove to a challenging party that the argument may be defended against any attack?

b) when the given argument *cannot* be defended against all possible attacks, how many rounds *must* it take to convince putative defenders that their position is untenable?

[*] Technical Report: Dept. of Comp. Sci., Univ. of Liverpool, Oct. 2001

We obtain a precise characterisation answering (a) (Theorem 4, below). In the case of (b), by developing a construction first presented in [16], the question is related to the widely studied issue of Proof Complexity. Specifically, we demonstrate that by representing an unsatisfiable CNF-formula, φ, as an argument system the dispute protocol defines a *proof calculus* that may be employed to show $\neg\varphi$ is a *propositional tautology*. Thus, we obtain a partial answer to (b) (in Theorem 5) by establishing that when interpreted as a calculus for Propositional Logic, the TPI-dispute protocol is 'not very powerful': formally we show that it may be efficiently simulated by a Gentzen system in which the CUT inference rule is not available.

In the remainder of this section we review the Argument System formalism from [15] and formally develop the argument game TPI-dispute, originally outlined in [34]. In Section 2 some illustrative examples of how disputes evolve in this protocol are presented. As we have already noted, Section 3 presents the core technical contribution, while Section 4 discusses some issues arising from our results and presents some directions for further work. Conclusions are given in Section 5.

1.1. ARGUMENT SYSTEMS

Argument Systems as a mechanism for studying formalisations of reasoning, acceptability, and defeasibility were introduced by Dung [15] and have since received considerable attention with respect to their use in non-classical logics, e.g. [8, 13, 14]. The basic definition below is derived from that given in [15].

DEFINITION 1. *An* argument system *is a pair $\mathcal{H} = \langle \mathcal{X}, \mathcal{A} \rangle$, in which $\mathcal{X}$ is a set of* arguments *and $\mathcal{A} \subset \mathcal{X} \times \mathcal{X}$ is the* attack relationship *for $\mathcal{H}$. Unless otherwise stated, $\mathcal{X}$ is assumed to be* finite*, and $\mathcal{A}$ comprises a set of ordered pairs of* distinct *arguments. A pair $\langle x, y \rangle \in \mathcal{A}$ is referred to as 'x* is attacked by *y' or 'y* attacks *(or* is an attacker of*) x'. The* range *of an argument x – denoted range(x) – is the set of arguments that are attacked by x; the* range of a set *of arguments S, is the union over all x in S of range(x).*

For R, S subsets of $\mathcal{X}$ in $\mathcal{H}(\langle \mathcal{X}, \mathcal{A} \rangle)$, we say that

a) $s \in S$ is attacked *by R if there is some $r \in R$ such that $\langle r, s \rangle \in \mathcal{A}$.*

b) $x \in \mathcal{X}$ is acceptable with respect to *S if for every $y \in \mathcal{X}$ that attacks x there is some $z \in S$ that attacks y.*

c) S is conflict-free *if no argument in S is attacked by any other argument in S.*

d) A conflict-free set S is admissible *if every argument in S is acceptable with respect to S.*

e) S is a preferred extension *if it is a maximal (with respect to $\subseteq$) admissible set.*

f) S is a stable extension *if S is conflict free and every argument $y \notin S$ is attacked by S.*

Our disqualification of 'self-attack' – $\langle x, x\rangle$ – is made purely to conform with standard graph-theoretic conventions. Although a similar constraint is not imposed in Dung's original framework, there is no significant effect on the class of systems with which this article is concerned: argument systems that contain an attack $\langle x, x\rangle$ can be simulated by creating two new arguments $\{x_1, x_2\}$ and replacing $\langle x, x\rangle$ with $\{\langle x, x_1\rangle, \langle x_1, x_2\rangle, \langle x_2, x\rangle\}$ so yielding an argument system in the sense of Defn. 1.

DEFINITION 2. *The decision problem* Credulous Acceptance *(*CA*) takes as an instance: an argument system $\mathcal{H} = \langle \mathcal{X}, \mathcal{A}\rangle$ and an argument $x \in \mathcal{X}$. The result true is returned if and only if* at least one *preferred extension S of $\mathcal{X}$ contains x. If* CA$\langle \mathcal{H}, x\rangle)$ *holds then x is said to be* credulously accepted *in $\mathcal{H}$.*

The decision problem Sceptical Acceptance *(*SA*) takes as an instance: an argument system $\mathcal{H} = \langle \mathcal{X}, \mathcal{A}\rangle$ and an argument $x \in \mathcal{X}$. The result true is returned if and only if* every *preferred extension S in $\mathcal{X}$ contains x. If* SA$(\langle \mathcal{H}, x\rangle)$ *holds x is said to be* sceptically accepted *in $\mathcal{H}$.*

While some argument systems may not have any stable extension, it is always the case that *some* preferred extension is present: the reason being that the empty set is always admissible.

1.2. ARGUMENT GAMES AND TPI-DISPUTES

A widely studied concept that has received some attention in the context of argument systems is that of employing *argument games* both as models of dialectical discourse and as a basis for a formal proof theory. The form of such games involves a sequence of interactions between two protagonists – hereafter referred to as the *Defender* (*D*) and *Challenger* (*C*) – wherein the Defender attempts to establish a particular argument in the face of counterarguments advanced by the Challenger, see e.g. [10, 21, 23, 26, 33]. In [34] descriptions of games – *Two Party Immediate Response Disputes* (TPI-disputes) – are presented for Credulous and Sceptical Argument within the framework considered in the present article. We consider a rather more tightly specified definition of TPI-disputes. Our main point of interest concerns the fact that whilst such games always terminate for finitely specified systems we wish to address how many steps (as a function of $|\mathcal{X}|$) some disputes may take.

We begin by developing the idea of TPI-disputes, using as a basis the informal schema of [34]. In informal terms, a TPI-dispute starts from a named argument, x in a given argument system $\mathcal{H}$. The Defender attempts to construct an admissible set containing x (for the *Credulous Game*) or to show that no attack on x can be in an admissible set[1] for the (*Sceptical Game*). The Challenger's aim is to prevent successful construction. The game proceeds by the players alternately presenting arguments within $\mathcal{H}$ that attack the previous arguments proposed by the other player. The concept of *immediate response* concerns the requirement in the game for both players to identify arguments that attack the most recent argument put forward by the opponent. A number of examples given in [34] indicate that both players must have the capability of 'back-tracking', e.g. if the line of attack followed by the Challenger fails, it must be possible to adopt a different attack on some previous argument.

We can view the progress of such disputes as a sequence of *directed trees* each of which is constructed by a *depth-first* expansion, the root of each tree being the argument x at the heart of the dispute. In this way the game is characterised by the *moves* through which a tree is expanded and the rules which force back-tracking by either party.

1.2.1. *A model of* TPI*-disputes*

DEFINITION 3. *Let $\mathcal{H}(\langle \mathcal{X}, \mathcal{A} \rangle)$ be an argument system and x an argument in $\mathcal{X}$. A* dispute tree for *x* in *$\mathcal{H}$, $T_x^{\mathcal{H}}$, is a tree whose vertices are a subset of $\mathcal{X}$ and whose root is x. The edges of a dispute tree are directed* from *vertices* to *their parent vertex. If t is a leaf vertex in $T_x^{\mathcal{H}}$ the path*

$$t = v_k \to v_{k-1} \to \cdots \to v_2 \to v_1 \to v_0 = x$$

is called a dispute line.

*A dispute line (*to *v) is a* failing attack on *x if the number of vertices on the path from v up to (and including) x is* odd. *A dispute line is a* failing defence of *x if this number is* even.

A vertex, v, is open *in $T_x^{\mathcal{H}}$ if there is an argument, w in $\mathcal{X}$, which attacks v and is* 'available' *(in a sense which is made precise below). If no such argument exists, v is* closed. *A dispute line is closed or open according to whether its leaf vertex is closed or open.*

Given a system $\mathcal{H}(\langle \mathcal{X}, \mathcal{A} \rangle)$ and $x \in \mathcal{X}$ a TPI-dispute consists of a sequence of moves

$$M = \langle \mu_1, \mu_2, \ldots, \mu_i, \ldots \rangle$$

Moves, μ are chosen from a finite *repertoire* of *move types*, some (or all) of which may not be available (depending on the current 'state' of a dispute). This *state* is represented after the k'th move ($k \geq 0$), by a tuple $\sigma_k =$

[1] This approach to establishing Sceptical Acceptance can not be applied to *all* argument systems, cf. Theorem 3, and Fig. 1 subsequently.

$\langle T_k, v_k, \Delta_k, \Gamma_k, P_k, Q_k \rangle$. Here

T_k	:	the dispute tree after k moves
v_k	:	the 'current' argument (vertex of) T_k
Δ_k	:	arguments available to D
Γ_k	:	arguments available to C
P_k	:	arguments proposed as a (subset) of some admissible set by D
Q_k	:	the set of *subsets* of arguments that C has *shown* not to be a subset of an admissible set

The initial state (σ_0) is $\langle \langle x \rangle, x, \Delta_0, \Gamma_0, P_0, Q_0 \rangle$ where

$$\begin{array}{rcl} \Delta_0 & = & \mathcal{X}/(\{x\} \cup \{y : \langle x, y \rangle \in \mathcal{A} \text{ or } \langle y, x \rangle \in \mathcal{A}\}) \\ \Gamma_0 & = & \mathcal{X}/(\{x\} \cup range(x)) \\ P_0 & = & \{x\} \\ Q_0 & = & \emptyset \end{array}$$

A dispute, $M = \langle \mu_1, \mu_2, \ldots, \mu_k \rangle$, is *active* if there is a legal move μ_{k+1} available to the current player, i.e. C if $k+1$ is odd, D otherwise. A dispute, M, is *terminated* if M is *not* active. For a terminated dispute, we use $|M|$ to denote the number of *moves* in M.

1.2.2. *The move repertoire*

It remains to describe the move repertoire, conditions determining applicability, and consequent changes to σ_{i-1} after performing a move μ_i.

The various implementations of argument games allow a variety of different moves. Some, such as [22], provide a small number of basic moves, intended to model disputes in a generic manner, while others, allow a larger number in order to attempt to reflect the moves made by the partcipants in particular kinds of dispute, e.g. [19] or to reflect particular notions of what constitutes an argument. For example Bench-Capon [6] models arguments as described by Toulmin [30]. Since our framework uses Dung's very abstract notion of argument [15], we do not need moves to reflect particular procedures or forms of argument, and so can use a rather small set of moves.

The repertoire of moves we allow comprises just,

$$\{\text{COUNTER}, \text{BACKUP}, \text{RETRACT}\}$$

The first move can be made by either player, whereas BACKUP is only employed by C, and RETRACT only by D. These two moves arise from the need to allow back-tracking. In the description that follows it should be remembered that *odd* indexed moves are made by the Challenger and *even* indexed moves by the Defender.

$\mu_k = \text{COUNTER}(y)$

Let $\sigma_{k-1} = \langle T_{k-1}, v_{k-1}, \Delta_{k-1}, \Gamma_{k-1}, P_{k-1}, Q_{k-1}\rangle$. If k is odd, μ_k is made by C, and COUNTER(y) can be applied only if $\langle y, v_{k-1}\rangle \in \mathcal{A}$ and $y \in \Gamma_{k-1}$, i.e y attacks the current argument (v_{k-1}) and is available. The new state, σ_k, is now

$$\begin{array}{rcl} T_k & := & T_{k-1} + \langle y, v_{k-1}\rangle \\ v_k & := & y \\ \Delta_k & := & \Delta_{k-1} \\ \Gamma_k & := & \Gamma_{k-1}/\{y\} \\ P_k & := & P_{k-1} \\ Q_k & := & Q_{k-1} \end{array}$$

If k is even, so that μ_k is made by D, then COUNTER(y) can be applied only if: $y \in \Delta_{k-1}$; $\langle y, v_{k-1}\rangle \in \mathcal{A}$; and for each set R in Q_{k-1}, R is not contained in $P_{k-1} \cup \{y\}$, i.e. D has available an argument y with which to attack v_{k-1} and, if y is added to the set of arguments that D is (currently) committed to then the resulting set has *not* been ruled inadmissible earlier.

The new state, σ_k is now

$$\begin{array}{rcl} T_k & := & T_{k-1} + \langle y, v_{k-1}\rangle \\ v_k & := & y \\ \Delta_k & := & \Delta_{k-1}/(\{y\} \cup \{z \in \Delta_{k-1} : \langle y, z\rangle \in \mathcal{A} \text{ or } \langle z, y\rangle \in \mathcal{A}\}) \\ \Gamma_k & := & \Gamma_{k-1}/(\{y\} \cup \mathit{range}(y)) \\ P_k & := & P_{k-1} \cup \{y\} \\ Q_k & := & Q_{k-1} \end{array}$$

The definition of Δ_k from Δ_{k-1} and y captures the fact that D (in attempting to form an admissible set) may not violate the requirement to be conflict free. The form taken by Γ_k indicates that in adding y to its (currently) accepted arguments, D now has a defence to all arguments in Γ_{k-1} that y attacks. It follows that there is no gain in these being available to C.

$\mu_k =$ BACKUP(j, y) (where j is *even* and $0 \leq j \leq k-3$)
The BACKUP move is only invoked by C and corresponds to the situation where C has no available attack with which to continue the current dispute line. The BACKUP move returns the dispute to the *most recent* point (σ_j) from which C can mount a fresh attack. Thus, if the (currently open) dispute line is,

$$L_{k-1} = \langle v_0 \leftarrow v_1 \leftarrow \cdots \leftarrow v_j \leftarrow v_{j+1} \leftarrow \cdots \leftarrow v_{k-3} \leftarrow v_{k-2} \leftarrow v_{k-1}\rangle$$

then

BC1. L_{k-1} is a closed failing attack, i.e. there are no arguments $z \in \Gamma_{k-1}$ for which $\langle z, v_{k-1}\rangle \in \mathcal{A}$.

BC2. For each r in the set $\{j+2, j+4, j+6, \ldots, k-3\}$ there are no arguments

$$z \in \Gamma_r/(\{v_r, v_{r+1}, v_{r+2}, \ldots, v_{k-2}\} \cup \mathit{range}(\{v_r, v_{r+2}, \ldots, v_{k-3}\})$$

for which $\langle z, v_r \rangle \in \mathcal{A}$.

BC3. The parameters j and y specified in the move BACKUP(j, y) are such that

$$y \in \Gamma_j / (\{v_j, v_{j+1}, v_{j+2}, \ldots, v_{k-2}\} \cup \mathit{range}(\{v_j, v_{j+2}, \ldots, v_{k-3}\})$$

and $\langle y, v_j \rangle \in \mathcal{A}$.

In summary, the conditions for the move BACKUP(j, y) to be applicable are: C cannot continue the current dispute line since there is no argument in C's arsenal that can be used to attack the last argument proposed by D (BC1); C cannot mount a *new* line of attack on any argument put forward by D in the set $\{v_{j+2}, v_{j+4}, \ldots, v_{k-3}\}$ (BC2); C, by using y, can launch a different attack on v_j (BC3).

The new state σ_k effected by the move BACKUP(j, y) is given by:

$$\begin{array}{lcl} T_k & := & T_{k-1} + \langle y, v_j \rangle \\ v_k & := & y \\ \Delta_k & := & \Delta_{k-1} \\ \Gamma_k & := & \Gamma_j / (\{y, v_{j+1}, v_{j+2}, \ldots, v_{k-1}\} \cup \ \mathit{range}(\{v_{j+2}, v_{j+4}, \ldots, v_{k-1}\})) \\ P_k & := & P_{k-1} \\ Q_k & := & Q_{k-1} \end{array}$$

Note that Δ_k does *not* revert to its content at the 'backup' position Δ_j: D has 'committed' to defending these in order to force C to adopt a new line of dispute. Secondly, the set, Γ_k, of available arguments for C, has all of the arguments advanced in progressing from v_{j+1} to v_{k-3} removed (rather than simply the 'old' attack v_{j+1} and the 'new' attack y on v_j): since D has already established a suitable line of defence to each of these, their only utility to the challenger would be in *prolonging* a dispute, rather than winning it.

$\mu_k =$ RETRACT

The RETRACT move is only made by D. Suppose

$$\sigma_{k-1} = \langle T_{k-1}, v_{k-1}, \Delta_{k-1}, \Gamma_{k-1}, P_{k-1}, Q_{k-1} \rangle$$

is the current state (so that $k-1$ is odd). For RETRACT to be applicable D must have no available attack on v_{k-1} *and* $P_{k-1} \neq \{x\}$. In this case, the Challenger has succeeded in showing that the set P_{k-1} cannot be extended to form an admissible set. Thus the only option available to the Defender is to try constructing a new admissible set containing x. Formally, the next state σ_k is given as

$$\begin{array}{lcl} T_k & := & \langle x \rangle \\ v_k & := & x \\ \Delta_k & := & \Delta_0 \\ \Gamma_k & := & \Gamma_0 \\ P_k & := & P_0 \\ Q_k & := & Q_{k-1} \cup \{P_{k-1}\} \end{array}$$

1.2.3. *Discussion*

The main point that should be noted is the asymmetry concerning BACKUP and RETRACT. Firstly, BACKUP may be seen as the Challenger invoking a *new line of attack* within the *same* dispute tree. On the other hand, RETRACT represents the dispute over x being *started again*, this time, however, with the knowledge that some lines of defence are not available, i.e. those that would result in a 'known' inadmissible set being constructed. Of course, as will be shown later, if x is *credulously accepted* then D, employing 'best play' will never need to make a retraction. In defining the game rules, however, we cannot assume that D will play 'intelligently' and thus may, inadvertently, call upon arguments that are eventually exposed as collectively indefensible. It may be observed that the position from which the dispute is resumed (following a retraction) is the *opening* dispute tree: while, in principle, one could define the next dispute tree to result from some variant of the current one, such an approach affords no significant gain.

1.2.4. *Credulous and Sceptical Games*

DEFINITION 4. *Let $M_{\langle\mathcal{H},x\rangle} = \langle\mu_1, \mu_2, \ldots, \mu_k\rangle$ be a* terminated TPI-*dispute over an argument x in the argument system $\mathcal{H}$. $M_{\langle\mathcal{H},x\rangle}$ is a* successful (credulous) defence of *x if k is* even*, and a* successful rebuttal of *x if k is* odd*.*

The following result reformulates Proposition 1 of [34] in terms of the formal framework introduced above.

THEOREM 1.

$$\mathrm{CA}(\mathcal{H}, x) \Leftrightarrow (\exists M_{\langle\mathcal{H},x\rangle} : M_{\langle\mathcal{H},x\rangle} \textit{ is a successful defence of } x)$$

Proof. First suppose that $\mathrm{CA}(\mathcal{H}(\langle\mathcal{X}, \mathcal{A}\rangle, x)$ holds, i.e. that x is credulously accepted in $\mathcal{H}$. Consider any preferred extension, S_x, of $\mathcal{H}$ containing x. It is certainly the case that using *only* the arguments in S_x, D can always COUNTER attacks available to C (recall that in replying to $\mathrm{COUNTER}(y)$ from C the response $\mathrm{COUNTER}(z)$ will remove from C's arsenal of attacks any argument in this attacked by z). Furthermore, D never has to invoke the RETRACT move. It follows that such a dispute will eventually terminate with C having no further move, i.e. as a successful defence of x.

Conversely, suppose that $M_{\langle\mathcal{H},x\rangle}$ is a successful defence of x. Consider the set P_k pertaining after μ_k the final move of the dispute. It is certainly the case that $x \in P_k$ (since this holds throughout the dispute). In addition, P_k is conflict-free (since Δ_j never makes available to D, arguments that conflict with those in P_j). Finally, since C has no move available, every attack on arguments $y \in P_k$ must have been countered, i.e. is defended by some $p \in P_k$. The three properties just identified establish that P_k is an admissible set containing x, hence x is credulously accepted. □

THEOREM 2. *For all* TPI*-dispute instances,* $\langle \mathcal{H}, x \rangle$ *either* all *terminated* $M_{\langle \mathcal{H},x \rangle}$ *are successful defences of x or all are successful rebuttals.*

Proof. Suppose the contrary and

$$M^{(1)} = \langle \mu_1^1, \mu_2^1, \ldots, \mu_m^1 \rangle \text{ with } \sigma_m^1 = \langle T_m^1, v_m^1, \Delta_m^1, \Gamma_m^1, P_m^1, Q_m^1 \rangle$$

$$M^{(2)} = \langle \mu_1^2, \mu_2^2, \ldots, \mu_n^2 \rangle \text{ with } \sigma_n^2 = \langle T_n^2, v_n^2, \Delta_n^2, \Gamma_n^2, P_n^2, Q_n^2 \rangle$$

are different TPI-disputes with $M^{(1)}$ a successful defence of x and $M^{(2)}$ a successful rebuttal of x within $\mathcal{H}$. Since $M^{(1)}$ is a successful defence, the subset P_m^1 is an admissible set (containing x). If $M^{(2)}$ is a successful rebuttal of x, then D must reach the point where no RETRACT move is applicable. Consider the admissible set, P_m^1 found by $M^{(1)}$ and the first move t at which some $Q \subseteq P_m^1$ is added to Q_{t-1}^2. It must be the case that $\mu_t^2 = \text{RETRACT}$ (or $t = n+1$) and that D has no available defence with which to counter v_{t-1}^2. Now we derive a contradiction: v_{t-1}^2 attacks $y \in P_{t-2}^2 = Q \subseteq P_m^1$ and the progress of $M^{(2)}$ has left no counter attack on v_{t-1}^2 available to D. On the other hand, such a defence (z, say) *is* present in P_m^1 since it is an admissible set and z would only be unavailable if it attacked or was attacked by Q, contradicting the fact that P_m^1 (of which Q is a subset) must be conflict-free. □

DEFINITION 5. *For an argument system* $\mathcal{H}(\langle \mathcal{X}, \mathcal{A} \rangle)$ *and* $x \in \mathcal{X}$*, the x-*augmented system*,* $\mathcal{H}_x$ *is the system formed by adding a new argument* $\{x_a\}$ *to* $\mathcal{X}$ *together with attack* $\{\langle x, x_a \rangle\}$*.*

The following reformulates Proposition 2 of [34].

THEOREM 3. *Let* $\mathcal{H}$ *be an argument system in which* every *preferred extension is also a stable extension and let x be an argument in* $\mathcal{H}$*. The argument x is* sceptically accepted *in* $\mathcal{H}$ *if and only if, there is a dispute, M, providing a successful rebuttal of* x_a *in the x-augmented system* $\mathcal{H}_x$*.*

Proof. Let $\mathcal{H}$ be an argument system in which every preferred extension is stable. First suppose that x is sceptically accepted in $\mathcal{H}$, the first part of the theorem will follow (via Thm. 1) by showing that x_a is not credulously accepted in the x-augmented system. Suppose the contrary and that $S_a \subset \mathcal{X} \cup \{x_a\}$ is a preferred extension in $\mathcal{H}_x$ that contains x_a. The set S_a cannot contain x, and must contain at least one attacker of x. The set, $S_a/\{x_a\}$, however, is an admissible set in $\mathcal{H}$ and cannot be developed to a preferred extension containing x. This contradicts the premise that x is sceptically accepted in $\mathcal{H}$.

Conversely, suppose that x_a is not credulously accepted in the x-augmented system $\mathcal{H}_x$. Consider any preferred extension S of $\mathcal{H}$. Suppose $x \notin S$. Since S is a stable extension, there is some attacker, y, of x, in S and since y attacks

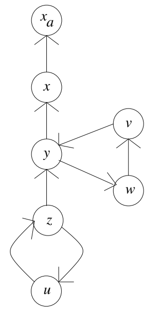


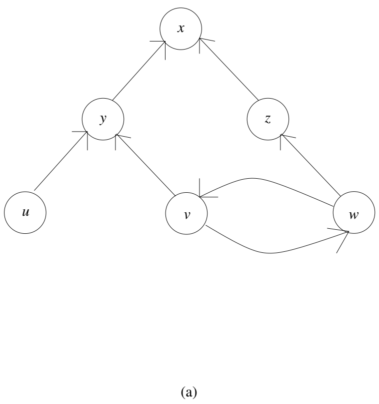
Figure 1. x-augmented system with x_a not credulously accepted nor x sceptically accepted.

x which is the only attack on x_a in the x-augmented system, we deduce that $S \cup \{x_a\}$ would form a preferred extension in $\mathcal{H}_x$ contradicting the premise that x_a is not credulously accepted. ▫

The example in Fig. 1 is adapted from [34], and shows that the stability condition is needed. In this example of an x-augmented system, x_a is not in any preferred extension since there is no defence to the attack by x (y is inadmissible since it is effectively self-attacking). Within the original system, however, x is *not* sceptically accepted: there are two preferred extensions – $\{x, z\}$ and $\{u\}$ – the latter containing neither x nor its attacker y.

2. Examples

In order to clarify how particular disputes develop we give two examples based on the argument systems, shown in Fig. 2.



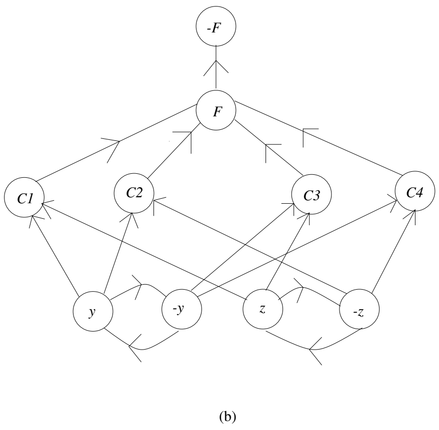


Figure 2. Two Example Argument Systems

It may be observed that the system in Fig. 2(b) can be interpreted as a representation of the tautology,

$$\neg F(y,z) = \neg((y \vee z) \wedge (y \vee \neg z) \wedge (\neg y \vee z) \wedge (\neg y \vee \neg z)) \qquad (1)$$

and so serves to illustrate dispute progression for proving *credulous acceptance* of the argument $\neg F$ *and sceptical acceptance* of the same argument, i.e. that the argument F in this system is *not* credulously accepted.

For Fig. 2(a) one possible TPI-dispute over x (in which we abbreviate COUNTER, BACKUP, and RETRACT to C,B,R) is,

k	μ_k	v_k	Δ_k	Γ_k	P_k	Q_k
0	–	x	$\{u,v,w\}$	$\{y,z,u,v,w\}$	$\{x\}$	$\emptyset$
1	$\mathrm{C}(y)$	y	$\{u,v,w\}$	$\{z,u,v,w\}$	$\{x\}$	$\emptyset$
2	$\mathrm{C}(v)$	v	$\{u\}$	$\{z,u\}$	$\{x,v\}$	$\emptyset$
3	$\mathrm{B}(0,z)$	z	$\{u\}$	$\{u\}$	$\{x,v\}$	$\emptyset$
4	R	x	$\{u,v,w\}$	$\{y,z,u,v,w\}$	$\{x\}$	$\{x,v\}$
5	$\mathrm{C}(y)$	y	$\{u,v,w\}$	$\{z,u,v,w\}$	$\{x\}$	$\{x,v\}$
6	$\mathrm{C}(u)$	u	$\{v,w\}$	$\{z,v,w\}$	$\{x,u\}$	$\{x,v\}$
7	$\mathrm{B}(4,z)$	z	$\{v,w\}$	$\{v,w\}$	$\{x,u\}$	$\{x,v\}$
8	$\mathrm{C}(w)$	w	$\emptyset$	$\emptyset$	$\{x,u,w\}$	$\{x,v\}$

(2)

It may be observed that D, at μ_2, makes an 'incorrect' move in attacking y using v (instead of u) thus removing w from the set of available arguments and allowing C to force a retraction by attacking x with z. Of course, D could have shortened the length of the dispute by playing $\mathrm{COUNTER}(u)$ as the second move. As we noted earlier, the intention is to define the protocol for disputes in such a way that even if D advances what turn out to be ill-advised counter-attacks, this does not result in the game being lost since there are opportunities to correct. For Fig. 2(a) there are exactly three possible TPI-disputes over x: two in which C first counter-attacks with y, and one in which the initial counter-attack is using z.

As a final illustration we give an example of a dispute establishing sceptical acceptance of $-F$ in the system of Fig. 2(b). It is not difficult to see that this follows by showing that F is not credulously accepted, so the description is given in terms of a successful rebuttal of F;

k	μ_k	v_k	P_k	Q_k
0	–	F	$\{F\}$	$\emptyset$
1	$\textsc{c}(C1)$	$C1$	$\{F\}$	$\emptyset$
2	$\textsc{c}(y)$	y	$\{F,y\}$	$\emptyset$
3	$\textsc{b}(0,C3)$	$C3$	$\{F,y\}$	$\emptyset$
4	$\textsc{c}(z)$	z	$\{F,y,z\}$	$\emptyset$
5	$\textsc{b}(0,C4)$	$C4$	$\{F,y,z\}$	$\emptyset$
6	$\textsc{r}$	F	$\{F\}$	$\{\{F,y,z\}\}$
7	$\textsc{c}(C1)$	$C1$	$\{F\}$	$\{\{F,y,z\}\}$
8	$\textsc{c}(z)$	z	$\{F,z\}$	$\{\{F,y,z\}\}$
9	$\textsc{b}(6,C2)$	$C2$	$\{F,z\}$	$\{\{F,y,z\}\}$
10	$\textsc{r}$	F	$\{F\}$	$\{\{F,y,z\},\{F,z\}\}$
11	$\textsc{c}(C1)$	$C1$	$\{F\}$	$\{\{F,y,z\},\{F,z\}\}$
12	$\textsc{c}(y)$	y	$\{F,y\}$	$\{\{F,y,z\},\{F,z\}\}$
13	$\textsc{b}(10,C3)$	$C3$	$\{F,y\}$	$\{\{F,y,z\},\{F,z\}\}$
14	$\textsc{r}$	F	$\{F\}$	$\{\{F,y,z\},\{F,z\},\{F,y\}\}$
15	$\textsc{c}(C1)$	$C1$	$\{F\}$	$\{\{F,y,z\},\{F,z\},\{F,y\}\}$

(3)

and now, D cannot counter-attack $C1$ without constructing an already shown to be inadmissible set nor RETRACT since $P_{15} = \{F\}$.

3. Complexity of Argument Games

The preceding sections have largely been concerned with a rigorous formulation of the concept of TPI-dispute as first outlined in [34]. The principal aim of the present paper, however, is to consider the following questions.

Question 1. Given a TPI-dispute instance – $\langle \mathcal{H}, x\rangle$ – such that x *is* credulously accepted in $\mathcal{H}$, how many moves are required (in the worst-case) in a dispute M defining a *successful defence* of x?

Question 2. Given a TPI-dispute instance – $\langle \mathcal{H}, x\rangle$ – such that x *is not* credulously accepted in $\mathcal{H}$, how many moves are necessary (in the best case) for a dispute M establishing a *successful rebuttal* of x?

In view of Theorem 3, Question 2, is of interest with respect to the number of moves required to establish *sceptical acceptance* of an argument.

In order to make these precise, we introduce the idea of *Dispute Complexity*. Given an instance of a TPI-dispute, $\langle \mathcal{H}, x\rangle$, its *dispute complexity*, denoted $\delta(\mathcal{H},x)$ is,

$$\delta(\mathcal{H},x) = \min_{M\,:\,M \text{ is a terminated dispute over } x \text{ in } \mathcal{H}} |M|$$

Question 1 turns out to have a relatively straightforward characterisation using the following idea.

DEFINITION 6. *Let $\mathcal{H}(\langle\mathcal{X},\mathcal{A}\rangle)$ be an argument system and $x \in \mathcal{X}$ an argument that is credulously accepted in $\mathcal{H}$. The* rank *of x* in $\mathcal{H}$, *denoted $\rho(\mathcal{H},x)$, is defined by*

$$\min_{S\subseteq\mathcal{X}/\{x\}\ :\ S\cup\{x\}\ \textit{is admissible in}\ \mathcal{H}} |S|$$

THEOREM 4. *For any* TPI*-dispute instance – $\langle\mathcal{H},x\rangle$ – in which x is credulously accepted in $\mathcal{H}$,*

$$\delta(\mathcal{H},x) = 2\rho(\mathcal{H},x)$$

Proof. To see that $\delta(\mathcal{H},x) \leq 2\rho(\mathcal{H},x)$, consider the subset S of $\mathcal{X}$ that attains the value $\rho(\mathcal{H},x)$. By an argument similar to that in the proof of Theorem 1, x can be defended in a TPI-dispute, with D employing only arguments in S. Adopting such a strategy, D never needs to invoke the RETRACT move. The size of the set, P, to which D is committed increases by one with each move made by D as more members of S are added. It follows that, since S is admissible, the Challenger will have no further moves open once D has committed to every argument in S. To complete the proof we show that $\delta(\mathcal{H},x) \geq 2\rho(\mathcal{H},x)$. Consider a TPI-dispute, M, that attains $\delta(\mathcal{H},x)$ and the dispute tree, $T_{|M|}$ that is active when the Challenger admits defeat. Certainly, $|M|$ must be at least twice the number of arguments in $T_{|M|}$ (excluding x). The arguments to which D is committed after the $|M|$'th move must define an admissible set (otherwise C could continue the dispute by finding an appropriate y attacking some member of $P_{|M|}$). It follows that $|P_{|M|}/\{x\}| \geq \rho(\mathcal{H},x)$ and thence $\delta(\mathcal{H},x) \geq 2\rho(\mathcal{H},x)$ as required. □

Theorem 4, in its characterisation of the answer to our first question, can be interpreted in the following way: *if* an argument x is credulously accepted in the system $\mathcal{H}$ then there is a 'short proof' of this, i.e. using the TPI-dispute that achieves $\delta(\mathcal{H},x)$ moves. It is important to note that this does not imply that *deciding* if such a proof *exists* can be accomplished efficiently: given the results of [16] (wherein CA is shown to be NP-complete) it seems unlikely that such a decision method could be found.

For the remainder of this paper we are concerned the second question raised. As with the view proposed in the preceding paragraph, we can interpret results concerning this question in terms of properties of the '*size*' of '*proofs*' that an argument is not credulously accepted. The decision problem CA being NP-complete, indicates that such proofs are concerned with a CO-NP-complete problem. While all NP-complete problems are such that positive instances of these have concise proofs that they *are* positive instances (this being one of the defining characteristics of the class NP as a whole) it is suspected that *no* CO-NP-complete problem has this property. In other words, we have the following (long-standing) conjecture: if L is a CO-NP-complete problem, then there are (infinitely many) instances, x of L, for which $L(x)$ is true but the 'shortest proof' of this is of length superpolynomial in the number of bits needed to encode x.[2]

The discussion above suggests that (assuming NP $\neq$ CO-NP) there must be infinitely many instances $\langle \mathcal{H}, x\rangle$ for which x is *not* credulously accepted in $\mathcal{H}$ *and* for which $\delta(\mathcal{H}, x)$ – the dispute complexity of the instance – is superpolynomial in $|\mathcal{X}|$, the number of arguments in the system. Our goal in the remainder of this paper is to establish the existence of a sequence of TPI-dispute instances – $\langle \mathcal{H}_N, x\rangle$ – having N arguments, x not credulously accepted in $\mathcal{H}_N$, and with the number of moves in any terminated TPI-dispute being exponential in N. Of course, since these bounds apply *only* to our specific formalisation, this raises the question of defining 'more powerful' dispute protocols.

3.1. Propositional Tautologies and Argument Systems

The proof that CA is NP-complete given in [16] is effected through a straightforward reduction from 3-SAT, it being observed in the course of the proof that it extends easily to CNF-SAT, i.e. without the restriction of three literals per clause. The class of argument systems that result via this translation of CNF formulae turn out to be central to the analysis of dispute complexity, we therefore review the details of this in,

DEFINITION 7. *Given,*

$$\varphi(Z_n) = \bigwedge_{i=1}^{m} C_i \; = \; \bigwedge_{i=1}^{m} \left(\bigvee_{j=1}^{k_i} y_{i,j} \right)$$

a propositional formula in CNF *comprising m clauses – C_i – the i'th containing exactly $k_i \geq 1$ distinct literals over the propositional variables $Z_n = \langle z_1, z_2, \ldots, z_n\rangle$, the argument system $\mathcal{H}_\varphi(\langle \mathcal{X}_\varphi, \mathcal{A}_\varphi\rangle)$ has $2n+m+1$ arguments*

$$\mathcal{X}_\varphi = \{\varphi\} \cup \{C_1, C_2, \ldots, C_m\} \cup \{z_1, \neg z_1, z_2, \neg z_2, \ldots, z_n, \neg z_n\}$$

and attack relationship $\mathcal{A}_\varphi$ in which,

1. $\forall C_i \; \langle C_i, \varphi\rangle \in \mathcal{A}_\varphi$.
2. $\forall z_j \; \langle z_j, \neg z_j\rangle \in \mathcal{A}_\varphi$ *and* $\langle \neg z_j, z_j\rangle \in \mathcal{A}_\varphi$.
3. $\langle z_j, C_i\rangle \in \mathcal{A}_\varphi$ *if z_j is a literal in the clause C_i.*

[2] In complexity-theoretic terms, this is the assertion that NP $\neq$ CO-NP. It is worth noting that if true, it implies P $\neq$ NP. The converse, however, is not (necessarily) true: in principle one might have NP $=$ CO-NP and P $\neq$ NP.

4. $\langle \neg z_j, C_i \rangle \in \mathcal{A}_\varphi$ *if* $\neg z_j$ *is a literal in the clause* C_i.

For convenience we will subsequently write $y \in C$ *rather than 'y is a literal in the clause C'.*

As a consequence of [16, Theorem 3] we have,

FACT 1. *The* CNF *formula* $\varphi(Z_n)$ *is satisfiable if and only if the argument* φ *is credulously accepted in the system* $\mathcal{H}_\varphi(\langle \mathcal{X}_\varphi, \mathcal{A}_\varphi \rangle)$.

Thus in attempting to derive lower bounds on dispute complexity for cases in which x is not credulously accepted in $\mathcal{H}$, we could focus on bounding $\delta(\mathcal{H}_\varphi, \varphi)$ for appropriate instances in which $\neg\varphi(Z_n)$ is a *tautology*, i.e. $\varphi(Z_n)$ is not satisfiable.

Our approach to establishing such lower bounds will be rather less direct than that of examining $\delta(\mathcal{H}_\varphi, \varphi)$ for a specific propositional tautology $\neg\varphi$. Instead, we shall show that the progression of a TPI-dispute over φ can be '*efficiently simulated*' within a specific Proof Calculus for Propositional Logic: since the calculus we employ is *known* to require exponentially long proofs to validate certain tautologies, it will then follow that $\delta(\mathcal{H}_\varphi, \varphi)$ for such φ must also be exponential (in the number of arguments defining $\mathcal{H}_\varphi$).

It is worth noting, at this point, that there is a rich corpus of research concerning the *length* of proofs in various proof systems. Results on the complexity of General Resolution date back to the seminal paper of Haken [20] in which this approach was shown to require exponential length proofs for tautologies corresponding to the combinatorial *Pigeon-Hole Principle*, with important subsequent work in, e.g [1, 3, 4, 27, etc.]. Excellent introductory surveys discussing progress involving proof complexity may be found in the articles by Pudlák[28] and Beame and Pitassi[5].

3.2. THE GENTZEN CALCULUS FOR PROPOSITIONAL LOGIC

The Proof Calculus around which our simulation is built is the *Gentzen* (or *Sequent*) *Calculus*, [18], with, however, one of its standard inference rules being unavailable.

In its most general (propositional) form, the Gentzen Calculus, prescribes rules for deriving *sequents* – $\Gamma \Rightarrow \Delta$ – where Γ, Δ are *sets* of propositional formulae (over a set of atomic propositional variables $\{x_1, x_2, x_3, \ldots\}$) built using some finite (complete) logical basis. A *proof* of the sequent $\Gamma \Rightarrow \Delta$, consists of a sequence of *derivation steps* each of which is either an axiom or follows by applying one of the rules to (at most) two previously derived sequents. In what follows we observe the convention of employing upper case Roman letters – $\{A, B, C, \ldots,\}$ – to denote propositional formulae, and upper case Greek letters – $\{\Gamma, \Delta, \ldots\}$ – to denote *sets* of such formulae. We use Γ, A to denote the set $\Gamma \cup \{A\}$.

DEFINITION 8. *(Gentzen Calculus for Propositional Formulae)*

Let $\mathcal{L}$ be the language of well-formed formulae using the basis $\{\wedge, \vee, \neg\}$ and propositional variables drawn from $\{z_1, z_2, z_3, \ldots\}$.

A sequent *is an expression the form $\Gamma \Rightarrow \Delta$ where Γ, Δ are (finite) subsets of $\mathcal{L}$, i.e. sets of well-formed formulae. For a sequent $S = \Gamma \Rightarrow \Delta$ we use* $\textsc{lhs}(S)$ *to denote Γ and, similarly,* $\textsc{rhs}(S)$ *to denote Δ. A* Gentzen System *is defined by a set $\mathcal{GS}$ of* axioms *and* inference rules. *A sequent $\Gamma \Rightarrow \Delta$ is* provable in the Gentzen System $\mathcal{GS}$ *(written $\vdash_{\mathcal{GS}} \Gamma \Rightarrow \Delta$), if there is a finite sequence of sequents,*

$$S_1, S_2, \ldots, S_{k-1}, S_k, S_{k+1}, \ldots, S_t$$

for which S_t is the sequent $\Gamma \Rightarrow \Delta$ and for all k ($1 \leq k \leq t$), the sequent S_k is either an axiom of $\mathcal{GS}$ or there are sequents S_i, S_j (with $i, j < k$) and an inference rule r of $\mathcal{GS}$ such that S_k may be inferred from S_i and S_j as a consequence of the rule r. The Proof Complexity *of a sequent S in the* Gentzen System *$\mathcal{GS}$ (denoted $\pi(S, \mathcal{GS})$ is defined for* provable *sequents, to be the least t such that S is derived by a sequence of t sequents.*[3]

We shall use a modification of the Gentzen system, $\mathcal{G}$ shown in Table I, wherein A and B are members of $\mathcal{L}$, and Γ, Δ, etc. subsets of $\mathcal{L}$.

Table I. The Gentzen System $\mathcal{G}$

Axioms			
	$\{A\} \Rightarrow \{A\}$		
Rules			
$(\theta \Rightarrow)$	$\dfrac{\Gamma \Rightarrow \Delta}{\Gamma, A \Rightarrow \Delta}$	$\dfrac{\Gamma \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, A}$	$(\Rightarrow \theta)$
$(\textsc{cut})$	$\dfrac{\Gamma \Rightarrow \Delta, A \;;\; \Gamma', A \Rightarrow \Delta'}{\Gamma \cup \Gamma' \Rightarrow \Delta \cup \Delta'}$		
$(\neg \Rightarrow)$	$\dfrac{\Gamma \Rightarrow \Delta, A}{\Gamma, \neg A \Rightarrow \Delta}$	$\dfrac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \Delta, \neg A}$	$(\Rightarrow \neg)$
$(\wedge \Rightarrow)$	$\dfrac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \wedge B \Rightarrow \Delta}$	$\dfrac{\Gamma \Rightarrow \Delta, A \;;\; \Gamma' \Rightarrow \Delta', B}{\Gamma \cup \Gamma' \Rightarrow \Delta \cup \Delta', A \wedge B}$	$(\Rightarrow \wedge)$
$(\vee \Rightarrow)$	$\dfrac{\Gamma, A \Rightarrow \Delta \;;\; \Gamma', B \Rightarrow \Delta'}{\Gamma \cup \Gamma', A \vee B \Rightarrow \Delta \cup \Delta'}$	$\dfrac{\Gamma \Rightarrow \Delta, A, B}{\Gamma \Rightarrow \Delta, A \vee B}$	$(\Rightarrow \vee)$

[3] We note that some authors choose to define proof complexity in terms of the total number of *symbol* occurrences over the derivation. For the class of propositional formulae we will be considering, the two measures are polynomially equivalent.

It may be observed that the Resolution Rule is, in fact, a special case of the CUT rule: if we consider clauses

$$P = x \vee \bigvee_{i=1}^{r} y_i \quad ; \quad Q = \neg x \vee \bigvee_{i=1}^{s} z_i$$

these are resolved (on x) to the clause

$$\bigvee_{i=1}^{r} y_i \vee \bigvee_{j=1}^{s} z_j$$

The clause P may be expressed as the sequent, $\{y_1, \ldots, y_r\} \Rightarrow \{\neg x\}$ and Q as $\{\neg x\} \Rightarrow \{\neg z_1, \ldots, \neg z_s\}$ whence the sequent $\{y_1, \ldots, y_r\} \Rightarrow \{\neg z_1, \ldots, \neg z_s\}$ follows from the CUT rule. For a more detailed comparison of General Resolution and Gentzen calculi the reader is referred to [2].

The Gentzen system that we will be considering is $\mathcal{G}/\text{CUT}$, i.e. that which allows all of the rules of the system $\mathcal{G}$ *except for* the CUT rule. We recall some standard results concerning the systems $\mathcal{G}$ and $\mathcal{G}/\text{CUT}$.

FACT 2. *(Gentzen[18]) The propositional formula $\mathcal{F} \in \mathcal{L}$ is a tautology if and only if* $\vdash_{\mathcal{G}} \emptyset \Rightarrow \{\mathcal{F}\}$.

FACT 3. *(The Gentzen Cut-Elimination Theorem,[18])*

$$\vdash_{\mathcal{G}} \emptyset \Rightarrow \{\mathcal{F}\} \text{ if and only if } \vdash_{\mathcal{G}/\text{CUT}} \emptyset \Rightarrow \{\mathcal{F}\}$$

Fact 3 establishes that the CUT rule is not needed in order to derive any provable sequent. Nonetheless, CUT turns out to be an extremely powerful operation:

FACT 4. *(Urquhart[31, 32]) There are (infinite) sequences of formulae $\langle \mathcal{F}_n \rangle$ in $\mathcal{L}$ for which:*

a. *$\mathcal{F}_n$ is a propositional tautology of n propositional variables.*

b. $\pi(\emptyset \Rightarrow \{\mathcal{F}_n\}, \mathcal{G}) = O(n^k)$ *(for $k \in \mathbf{N}$)*

c. $\pi(\emptyset \Rightarrow \{\mathcal{F}_n\}, \mathcal{G}/\text{CUT}) = \Omega\left(2^{n^{\epsilon}}\right)$ *(where $\epsilon > 0$)*

These constructions by Urquhart are explicit, i.e. a specific sequence $\langle \mathcal{F}_n \rangle$ is proved to have the properties stated in Fact 4.

We now state and prove the main theorem of this paper.

THEOREM 5. *Let*

$$\varphi(Z_n) = \bigwedge_{i=1}^{m} C_i \quad = \bigwedge_{i=1}^{m} \left(\bigvee_{j=1}^{k_i} y_{i,j} \right)$$

be any unsatisfiable CNF-*formula, $\mathcal{H}_\varphi$ be the argument system defined from $\varphi(Z_n)$ as given in Defn. 7, and S_φ the (provable) sequent,*

$$\emptyset \Rightarrow \bigcup_{i=1}^{m} \left\{ \neg \left(\bigvee_{j=1}^{k_i} y_{i,j} \right) \right\}.$$

Then,

$$\pi\left(S_\varphi, \mathcal{G}/\text{CUT}\right) \leq \delta(\mathcal{H}_\varphi, \varphi) + 2n + m \tag{4}$$

Less formally, Theorem 5 states that the length of the shortest proof of $\neg\varphi$ (φ being in CNF) being a tautology within the CUT-free Gentzen system cannot be 'much greater than' the number of moves needed to form a successful rebuttal φ in the argument system $\mathcal{H}_\varphi$.

Proof. (of Theorem 5) Let, φ, $\mathcal{H}_\varphi$, and S_φ be as described in the Theorem statement. Given *any* terminated TPI-dispute, M over φ in $\mathcal{H}_\varphi$ we describe how its progress may simulated in the Gentzen system $\mathcal{G}/\text{CUT}$. We first observe two important properties of the dispute M.

Firstly, M may be encoded as a sequence of *ordered* sets, R_i, (for which the term *retraction round* will subsequently be employed). Each of these takes the form

$$R_i = \langle D_1, y_1, D_2, y_2, \ldots, D_j, y_j, \ldots, D_q, y_q, F_i \rangle \tag{5}$$

where,

$$\begin{array}{rcl} \{D_1, D_2, \ldots, D_q, F_i\} & \subseteq & \{C_1, C_2, \ldots, C_m\} \\ y_j & \in & D_j \\ y_j & \notin & F_i \ \forall\, 1 \leq j \leq q \end{array}$$

In other words, R_i describes the alternation between clauses (D) used to attack φ and counterattacks (y) used to repel these attacks. The final attack by the clause F_i is the position at which the retraction of $\{\varphi, y_1, y_2, \ldots, y_q\}$ is forced. We observe that $|R_1|$ is the number of moves made in M prior to the *first* RETRACT move; and in general, $|R_i|$ is the number of moves between the retraction arising from R_{i-1} and the next such in M.

In the final move of M, the corresponding set R, contains just a single clause: i.e. that clause of φ upon which the Defender, by reason of the totality of earlier retractions, can mount no attack.

The second property of interest concerns the relationship between the literals defining a *retraction forcing clause*, F, and those used to defend against attacks on φ within the current dispute tree, i.e. the literals $\{y_1, y_2, \ldots, y_q\}$. The literals in F may be partitioned into two sets,

$$W = \{w_1, w_2, \ldots, w_r\} \ \ ; \ \ U = \{u_1, u_2, \ldots, u_s\} \tag{6}$$

wherein the literals in W cannot be used to attack F since for each $w \in W$, $\neg w \in \{y_1, y_2, \ldots, y_q\}$ and those in U are unavailable since for each $u \in U$, there is some subset V of $\{y_1, y_2, \ldots, y_q\}$ such that the Defender has retracted $\{\varphi, V, u\}$ in an earlier move.

With the two observations above, the idea underlying the proof may be described, informally, as efficiently deriving sequents that simulate the reasoning through which retractions are forced. More precisely, given

$$\langle R_1, R_2, \ldots, R_i, \ldots, R_t \rangle$$

the sequence of retraction rounds describing the dispute M, we construct a mapping $\beta : \{1, 2, \ldots, t\} \to N$ and sequents

$$\langle S_1, S_2, S_3, \ldots, S_p \rangle$$

for which

$$\begin{array}{lllll} \beta(i+1) & > & \beta(i) & \geq 1 & (1 \leq i < t) \\ \beta(t) & \leq & p & \leq \beta(t) + m & \end{array}$$

and

$$S_p = S_\varphi = \emptyset \Rightarrow \bigcup_{i=1}^{m} \left\{ \neg \left(\bigvee_{j=1}^{k_i} y_{i,j} \right) \right\}.$$

In general, the sequent $S_{\beta(i)}$ will express the fact that the Defender must retract the set $\{\varphi, y_1, y_2, \ldots, y_q\}$ in the i'th round, since this leaves no defence available to an attack by the clause F_i on φ.

To avoid a surfeit of subscripts, we use Y_i to denote the set $\{y_1, y_2, ..., y_q\}$ of literals defining R_i, with W_i and U_i being the partition of the retraction forcing clause, F_i, as described in (6) (obviously the *exact number* of literals in each of these will be dependent on which retraction round R_i is relevant).

When $U_i \neq \emptyset$, for each $u \in U_i$, $ret(u, Y_i)$ is a minimal (with respect to $\subseteq$) subset of Y_i for which the set of arguments $\{\varphi, u, ret(u, Y_i)\}$ has been the subject of an *earlier* retraction.[4] Finally, $index(u, Y_i)$ is,

$$index(u, Y_i) \; = \; \max \{k \leq \beta(i-1) \; : \text{LHS}(S_k) = ret(u, Y_i) \cup \{u\}\} \qquad (7)$$

[**Note:** That $index(u, Y_i)$ is well-defined will be clear from the remainder of the proof.]

The theorem will follow from the Claim below.

Claim. Given $\langle R_1, \ldots, R_t \rangle$ the sequence of retraction rounds defined by M, there is a mapping $\beta : \{1, 2, \ldots, t\} \to N$, with the following properties: $\beta(i+1) > \beta(i) > 1$; and, if the sequent, $S_{\beta(i)}$ is defined to be

$$S_{\beta(i)} = \{Y_i\} \Rightarrow \{\neg F_i\} \cup \bigcup_{u \in U_i} \text{RHS}(S_{index(u, Y_i)})$$

then

[4] An indefinite article is required here, since there may be more than one such subset, e.g. $\{\varphi, y_1, u\}$ and $\{\varphi, y_2, u\}$ could *both* have been retracted: the subsequent argument will show that in such cases, $ret(u, Y)$ can be chosen to be either $\{y_1\}$ or $\{y_2\}$.

a) $S_{\beta(i)}$ is well-defined, i.e. $index(u, Y_i)$ is defined for each $u \in U_i$.

b) $S_{\beta(i)}$ is provable in $\mathcal{G}/\text{CUT}$ with $\pi(S_{\beta(i)}, \mathcal{G}/\text{CUT}) \leq \beta(i)$.

Proof. First note that we may use the following derivations as the first $2n$ lines, prior to establishing $S_{\beta(1)}$. In consequence, $\beta(1) > 2n$.

Sequent	via	Line
$\{z_j\} \Rightarrow \{z_j\}$	Axiom	$2j-1$
$\{z_j, \neg z_j\} \Rightarrow \emptyset$	$(\neg \Rightarrow)$ and $2j-1$	$2j$

We complete the proof of the claim by induction on $i \geq 1$. The inductive base, $i = 1$, deals with the retraction enforced by R_1, i.e. we need to show that the sequent

$$S_{\beta(1)} = \{Y_1\} \Rightarrow \{\neg F_1\}$$

is derivable. Noting that R_1 represents the *first* occurrence of a RETRACT move by the Defender, the set U_1 must be empty, i.e. the retraction is forced because each literal that could be used to attack F_1 is unavailable by reason of Y_1 containing its negation. It follows that,

$$\begin{aligned} F_1 &= W_1 = \{w_1, w_2, \ldots, w_r\} \\ Y_1 &= \{\neg w_1, \neg w_2, \ldots, \neg w_r, y_{r+1}, y_{r+2}, \ldots, y_q\} \end{aligned}$$

Let T_k (for $1 \leq k \leq r$) be the sequent,

$$\{\neg w_1, \neg w_2, \ldots, \neg w_k\}, A_k \Rightarrow \emptyset \text{ where } A_k = \bigvee_{j=1}^{k} w_j$$

For $k = 1$, the sequent $T_1 = \{w_1, \neg w_1\} \Rightarrow \emptyset$ has already been derived. For $k > 1$, T_k is derived in one step from the sequent $\{w_k, \neg w_k\} \Rightarrow \emptyset$ and T_{k-1} by a single application of the rule $(\vee \Rightarrow)$. We deduce that,

$$\{\neg w_1, \neg w_2, \ldots, \neg w_k\}, F_1 \Rightarrow \emptyset$$

is derived in $k-1$ steps, and the required sequent - $S_{\beta(1)}$ - by a single application of $(\Rightarrow \neg)$ to T_r followed by $q - r$ applications of $(\theta \Rightarrow)$ in order to construct

$$\{\neg w_1, \ldots, \neg w_r, y_{r+1}, \ldots, y_q\} \Rightarrow \{\neg F_1\}$$

This gives the value of $\beta(1)$ as $2n + q$, where we note that μ_{2q+2} is the *first* RETRACT move occurring in M.

For the Inductive Step, we assume for all retraction rounds R_j with $1 \leq j < i$ that the following hold:

IH1) The value of $\beta(j)$ has been defined.

IH2) The sequent,

$$S_{\beta(j)} \;=\; \{Y_j\} \Rightarrow \{\neg F_j\} \cup \bigcup_{u \in U_j} \mathrm{RHS}(S_{index(u,Y_j)})$$

has been derived in $\mathcal{G}/\mathrm{CUT}$ after $\beta(j)$ steps.

To complete the inductive proof of the Claim, we 'simulate' the retraction round R_i and to this end it is necessary to,

C1) define a value of $\beta(i)$ which is greater than $\beta(i-1)$, and

C2) show that the sequent,

$$\{Y_i\} \Rightarrow \{\neg F_i\} \cup \bigcup_{u \in U_i} \mathrm{RHS}(S_{index(u,Y_i)})$$

is well-defined *and* derivable in a further $\beta(i) - \beta(i-1)$ steps.

Consider the retraction forcing clause, $F_i = \{W_i, U_i\}$, so that

$$Y_i = \{\neg w_1, \neg w_2, \ldots, \neg w_r, y_{r+1}, y_{r+2}, \ldots, y_q\}$$

If $U_i = \emptyset$, then with $\beta(i) = \beta(i-1) + q$, the sequent,

$$S_{\beta(i)} = \{Y_i\} \Rightarrow \{\neg F_i\}$$

is derivable in a further q steps using exactly the same approach as employed in the Inductive Base. Thus we may assume that U_i is non-empty with

$$U_i = \{u_1, u_2, \ldots, u_s\}$$

Recalling that $\langle W_i, U_i\rangle$ is a *partition* of F_i it is certainly the case that neither $\neg u \in Y_i$ nor $u \in Y_i$ (the latter holding since F_i was available to the Challenger with which to attack φ). This being so and u being unavailable to the Defender to attack F_i it follows that there has been a retraction round in which some subset of Y_i *together with* u (and φ) have been retracted. Therefore, some such subset of Y_i must satisfy the criteria defining $ret(u, Y_i)$ with respect to u. Suppose R_j is the round at which a committment to $\{\varphi, u, ret(u, Y_i)\}$ was retracted by the Defender. Clearly, $j < i$ and hence from the Inductive Hypothesis, the sequent, $S_{\beta(j)}$, with,

$$S_{\beta(j)} = \{u, ret(u, Y_i)\} \Rightarrow \Delta \text{ where } \emptyset \subset \Delta \subseteq \{\neg C_1, \ldots, \neg C_m\}$$

has been derived. As a result we deduce that for each $u \in U_i$, the value $index(u, Y_i)$ is defined and does not exceed $\beta(i-1)$. In summary, we have proven (via the Inductive Hypothesis) the existence of $s = |U_i|$ sequents,

$$\langle S_{i,1}, S_{i,2}, \ldots, S_{i,s}\rangle$$

for which

$$\text{LHS}(S_{i,k}) = ret(u_k, Y_i) \cup \{u_k\} \quad \text{and} \quad \text{RHS}(S_{i,k}) \subseteq \{\neg C_1, \neg C_2, \ldots, \neg C_m\}$$

We can now complete the derivation of the required sequent $S_{\beta(i)}$.

From $s - 1$ applications of $(\vee \Rightarrow)$ using $S_{i,1}, S_{i,2}, \ldots, S_{i,s}$ we obtain

$$S_{\beta(i-1)+s-1} = \left\{ \bigcup_{k=1}^{s} ret(u_k, Y_i) \right\}, \bigvee_{k=1}^{s} u_k \Rightarrow \left\{ \bigcup_{k=1}^{s} \text{RHS}(S_{i,k}) \right\}$$

A further r applications of $(\vee \Rightarrow)$ involving $S_{\beta(i-1)+s-1}$ and the sequents

$$\{w_k, \neg w_k\} \Rightarrow \emptyset$$

yields $S_{\beta(i-1)+r+s-1}$ as

$$\left\{ \bigcup_{k=1}^{s} ret(u_k, Y_i) \right\} \cup \left\{ \bigcup_{k=1}^{r} \neg w_k \right\}, \bigvee_{k=1}^{s} u_k \vee \bigvee_{k=1}^{r} w_k \Rightarrow \left\{ \bigcup_{k=1}^{s} \text{RHS}(S_{i,k}) \right\}$$

Recalling that,

$$F_i = \left(\bigvee_{k=1}^{s} u_k \vee \bigvee_{k=1}^{r} w_k \right)$$

a single application of $(\Rightarrow \neg)$ to $S_{\beta(i-1)+r+s-1}$ gives $S_{\beta(i-1)+r+s}$ as,

$$\left\{ \bigcup_{k=1}^{s} ret(u_k, Y_i) \right\} \cup \left\{ \bigcup_{k=1}^{r} \neg w_k \right\} \Rightarrow \left\{ \bigcup_{k=1}^{s} \text{RHS}(S_{i,k}) \right\}, \neg F_i$$

Finally, since it may be the case that

$$\left\{ \bigcup_{k=1}^{s} ret(u_k, Y_i) \right\} \cup \left\{ \bigcup_{k=1}^{r} \neg w_k \right\} \subset Y_i$$

(i.e. a *strict* subset of Y_i) a total of,

$$\left| Y_i \Big/ \left\{ \bigcup_{k=1}^{s} ret(u_k, Y_i) \cup \bigcup_{k=1}^{r} \neg w_k \right\} \right|$$

applications of $(\theta \Rightarrow)$ will give $S_{\beta(i)}$ as,

$$S_{\beta(i)} = \{Y_i\} \Rightarrow \{\neg F_i\} \cup \bigcup_{u \in U_i} \text{RHS}(S_{index(u,Y_i)})$$

where

$$\beta(i-1) + r + s \leq \beta(i) \leq \beta(i-1) + r + s + q \leq \beta(i-1) + 2q$$

Note that $2q = |R_i| - 1$ is the total number of moves occurring in M between the retraction round R_{i-1} and R_i. This completes the inductive proof of the Claim. □

To complete the proof of the Theorem we need only observe that the total number of steps required to derive S_φ is bounded above by $\beta(t) + m$.

[The additional m arises from the possibility that $S_{\beta(t)}$ may be of the form $\emptyset \Rightarrow \Delta$ with Δ a (non-empty) *strict* subset of

$$\{\neg C_1, \neg C_2, \ldots, \neg C_m\}$$

This could occur if some subset ψ of φ's clauses defined an unsatisfiable CNF-formula. In such cases $S_{\beta(t)}$ would not be *identical* to the sequent S_φ of the Theorem statement, however, at most m applications of $(\Rightarrow\ \theta)$ (adding the 'missing' $\neg C_i$ clauses) will suffice to derive S_φ from $S_{\beta(t)}$.]

From the analysis in the proof of the Claim it is clear that the values $\beta(i)$ satisfy:

$$\begin{array}{rcll} \beta(i) & \leq & \beta(i-1) + |R_i| & \text{when } 1 < i \leq t \\ \beta(1) & \leq & 2n + |R_1| & \end{array}$$

hence $\beta(t) \leq 2n + \sum_{i=1}^{t} |R_i| \leq 2n + |M|$.

Thus from any terminated TPI-dispute, M, over the unsatisfiable CNF-formula φ in the argument system $\mathcal{H}_\varphi$ we may construct a proof in $\mathcal{G}/\text{CUT}$ that $\neg\varphi$ is a tautology, i.e. of the sequent S_φ. Since this proof involves at most $|M| + 2n + m$ steps we conclude that

$$\pi(S_\varphi, \mathcal{G}/\text{CUT}) \ \leq\ \delta(\mathcal{H}_\varphi, \varphi) + 2n + m$$

as required. □

From Theorem 5 we get,

COROLLARY 1. *There are (infinite) sequences of argument system with arguments $x \in \mathcal{X}$ not credulously accepted but with the number of moves in any* TPI*-dispute establishing such exponential in $|\mathcal{X}|$.*

To conclude this Section, we illustrate how the example of Fig. 2(b) that resulted in the dispute given in (3) translates into a derivation of the required sequent following the proof in Theorem 5.

3.3. Example

Recall that Fig. 2(b) could be interpreted as the tautology

$$\neg F(y,z) = \neg((y \vee z) \wedge (y \vee \neg z) \wedge (\neg y \vee z) \wedge (\neg y \vee \neg z)) \tag{8}$$

From (3) using the encoding of retraction rounds described in the proof of Theorem 5

$$\begin{array}{rcl} R_1 & = & \langle (y \vee z), y, (\neg y \vee z), z, (\neg y \vee \neg z) \rangle \\ R_2 & = & \langle (y \vee z), z, (y \vee \neg z) \rangle \\ R_3 & = & \langle (y \vee z), y, (\neg y \vee z) \rangle \\ R_4 & = & \langle (y \vee z) \rangle \end{array} \tag{9}$$

The sequent we wish to derive is

$$\emptyset \Rightarrow \{\neg(y \vee z), \neg(y \vee \neg z), \neg(\neg y \vee z), \neg(\neg y \vee \neg z)\} \tag{10}$$

Following the mechanism in the Theorem, for R_1 we wish to derive

$$S_{\beta(1)} = \{y, z\} \Rightarrow \{\neg(\neg y \vee \neg z)\}$$

This is obtained by

Sequent	via	Line
$\{y\} \Rightarrow \{y\}$	Axiom	1
$\{y, \neg y\} \Rightarrow \emptyset$	1, $(\neg \Rightarrow)$	2
$\{z\} \Rightarrow \{z\}$	Axiom	3
$\{z, \neg z\} \Rightarrow \emptyset$	3, $(\neg \Rightarrow)$	4
$\{y, z\}, (\neg y \vee \neg z) \Rightarrow \emptyset$	$2, 4$, $(\vee \Rightarrow)$	5
$\{y, z\} \Rightarrow \{\neg(\neg y \vee \neg z)\}$	5, $(\Rightarrow \neg)$	6

Hence $\beta(1) = 6$.

For R_2 the sequent required is

$$S_{\beta(2)} = \{z\} \Rightarrow \{(\neg(y \vee \neg z), \neg(\neg y \vee \neg z)\}$$

where we use the fact that *ret*$(y, \{z\}) = \{z\}$, so that *index*$(y, \{z\}) = 6$.

Sequent	via	Line
$\{z\}, (y \vee \neg z) \Rightarrow \{\neg(\neg y \vee \neg z)\}$	$4, 6$, $(\vee \Rightarrow)$	7
$\{z\} \Rightarrow \{\neg(\neg y \vee \neg z), \neg(y \vee \neg z)\}$	7, $(\Rightarrow \neg)$	8

whence $\beta(2) = 8$. Notice that in deriving S_7, $\text{LHS}(S_4)$ is viewed as $\{z\}, \neg z$ and $\text{LHS}(S_6)$ as $\{z\}, y$, i.e. with $\Gamma = \Gamma' = \{z\}$, $A = \neg z$, and $B = y$ when the inference rule $(\vee \Rightarrow)$ of Table I is used.

For R_3 the sequent required is

$$S_{\beta(3)} = \{y\} \Rightarrow \{\neg(\neg y \vee z), \neg(y \vee \neg z), \neg(\neg y \vee \neg z)\}$$

where we use the fact that *ret*$(z, \{y\}) = \emptyset$, so that *index*$(z, \{y\}) = 8$.[5]

Sequent	via	Line
$\{y\}, (\neg y \vee z) \Rightarrow \{\neg(\neg y \vee \neg z), \neg(y \vee \neg z)\}$	$2, 8$, $(\vee \Rightarrow)$	9
$\{y\} \Rightarrow \{\neg(\neg y \vee \neg z), \neg(y \vee \neg z), \neg(\neg y \vee z)\}$	9, $(\Rightarrow \neg)$	10

[5] Were *ret*$(z, \{y\})$ not subject to a minimality condition, it could also be chosen as $\{y\}$, giving *index*$(z, \{y\}) = 6$. This choice would, in fact, still lead to a proof of the required final sequent. We also note the need for *index*$(z, \{y\})$ to be maximal since $\text{LHS}(S_3) = \{z\}$

giving $\beta(3) = 10$.

Finally for R_4 we have

$$\begin{array}{lll} \mathit{ret}(y,\emptyset) = \emptyset & \text{with} & \mathit{index}(y,\emptyset) = 10 \\ \mathit{ret}(z,\emptyset) = \emptyset & \text{with} & \mathit{index}(z,\emptyset) = 8 \end{array}$$

so that using S_8, S_{10} and $(\vee \Rightarrow)$ gives

$$S_{11} = \{(y \vee z)\} \Rightarrow \{\neg(\neg y \vee \neg z), \neg(y \vee \neg z), \neg(\neg y \vee z)\}$$

and with a single application of $(\Rightarrow \neg)$ to S_{11}, we derive the required sequent

$$S_{12} = \emptyset \Rightarrow \{\neg(y \vee z), \neg(y \vee \neg z), \neg(\neg y \vee z), \neg(\neg y \vee \neg z)\}$$

4. Discussion and Further Work

In this paper our primary goal has been to formalise the argument game (TPI-dispute) introduced in [34] and to analyse this in terms of one particular computational measure – dispute complexity. For what is technically the most interesting case – the length of dispute required to convince Defenders of an argument that their position is untenable – we have shown in Theorem 5 that applying this dispute regime to simple argument system representations of propositional tautologies occasions a form of proof calculus. This calculus is in one sense, however, extremely limited: any proof within it being capable of description by a comparable length proof in a CUT-free Gentzen System. Since examples are known of tautologies where allowing CUT admits exponentially shorter proofs[6], the protocol enforced by TPI-disputes when applied to certain propositional argument systems may take significantly longer to reach a conclusion than 'more powerful' deductive systems. We noted earlier, in describing the semantics of the RETRACT move that the position reverted to is the *initial* argument, rather than some 'intermediate' state of the dispute tree being developed. Among the reasons for favouring returning to the initial position, is that the length of disputes (as indicated by our simulation using CUT-free Gentzen Systems) does not, primarily, result from potentially repeating chains of defence which will ultimately fail: if the retraction mechanism were to revert to a 'sub-tree' of the dispute tree, cf. in a similar manner to that of the Challenger's BACKUP move, then this could be simulated from the initial argument just by repeating the relevant COUNTER and BACKUP moves. Since the size of any dispute tree can be at most the number of arguments within the system itself, a more sophisticated RETRACT semantics

[6] In fact, [32], shows $\mathcal{G}/\text{CUT}$ can be weaker than simple truth-tables proving worst-case lower bounds of $\Omega(n!)$ for the former as opposed to upper bounds of $n2^n$ for the latter.

could only shorten the length of a dispute by a polynomial factor – not reduce it exponentially.

Before dealing with some questions that are raised by the main result of this paper, it may be useful to place our concerns in the general context of argument systems, dialogue games, reasoning systems, etc. While the view of dialogue process as a 2-player game has been long established, e.g. MacKenzie's DC [24], interpretations of Toulmin's Argument Schema [30] as a game-based method [6], etc., the direction towards which such work has tended is in attempting formally to capture different types of dialogue process: e.g. [19] is, primarily, concerned with argument in a legal reasoning context. As a result there is a wealth of differing models of dialogue ranging from taxonomies of dialogue types as in Reed[29] and Walton and Krabbe[35] to frameworks modelling divers concepts of what 'winning' a dialogue game might mean, e.g [21]. Despite this variety of approaches, one unifying trend is that the central concern is primarily semantic, i.e. in defining the form(s) that games take, the rules and processes by which games evolve, the conditions under which games terminate, and in establishing degrees of soundness and completeness of the game capabilities. The question of how 'efficient' such processes might be, however, seems to have been largely neglected, with the exception of general complexity-theoretic classifications of Argumentation Frameworks within specific non-classical logics, e.g. [13, 14] or analyses of termination properties. Thus, little work is evident concerning more general contexts for the two questions which this paper has considered, i.e. with different protocols for the conduct of dialogues, different attack semantics, concepts of 'winning' other than credulous acceptance. If practical applications of dialectic and reasoning games are to be realised – as has become widely posited with the advent of autonomous agent systems – then measures analogous to our concept of dispute complexity may be of importance in evaluating implemented systems.

A rather different situation to that outlined in the preceding paragraph, pertains with respect to concepts of Proof Complexity, that we have used as the basis of our analysis of dispute complexity: Cook and Reckhow[12] introduced a formal mechanism for comparing the *complexity* of different proof calculi so that two 'different' systems are regarded as equipotent if a formal proof in one can be 'simulated' in the other with only a small increase in size. An important feature of this approach is that it can be developed to address questions concerning proof strategies for acceptance of instances in CO-NP-problems other than UNSAT, e.g. the Graph Stability Number calculus of Chvátal[11], or the Hajós Calculus for proving a graph has chromatic number greater than 3, [7, 25]. It is the case, however, that these analyses are effectively only dealing with Classical (Propositional) Logic, and such results as extend to non-classical Logics do so only by virtue of propositional logic being treatable as a sub-case, e.g. Haken[20] trivially applies to the

Resolution Calculus for Temporal Logic of [17] simply by expressing the relevant tautology without the use of any temporal operators, i.e. exactly as its propositional form.

We conclude by reviewing some directions for further research, that encompass both argument and dialogue game developments as well as extensions to the concept of dispute complexity.

Within the framework of [12] while it is known that the Gentzen System $\mathcal{G}/\mathrm{CUT}$ is weaker than both the system $\mathcal{G}$ and Propositional Proof systems employing General Resolution only, it is an open problem as to whether $\mathcal{G}$ and Resolution are equivalent, i.e. it has yet to be shown that, e.g. the Pigeon-Hole Principle tautologies require exponential length proofs in $\mathcal{G}$, however no (*efficient*) simulation of $\mathcal{G}$ by Resolution has been constructed. Theorem 5 establishes that using the TPI-dispute protocol as a vehicle for constructing proofs of propositional tautologies, $\neg\varphi$, affords a system which is (at best) as good as $\mathcal{G}/\mathrm{CUT}$, thus we might represent the respective power of various proof calculi for propositional tautologies informally as,

$$\mathcal{G} \;\geq\; \text{Resolution} \;>\; \mathcal{G}/\mathrm{CUT} \;\geq\; \mathrm{TPI} \qquad (11)$$

The situation depicted in (11) raises some interesting questions. Firstly, it may be noted that Theorem 5 operates in only 'one direction', that is we express the problem of proving a propositional formula to be a tautology as a problem of showing an argument is not credulously accepted in an argument system, thence relating a calculus for the latter to a calculus for the former. We have not considered, however, translations of argument systems *into* propositional formulae. Such translations may be of interest for a number of reasons: they may provide some insight into the formulation of different dispute protocols by building analogues of propositional proof calculi applied to the formula resulting from a translation; and, they may provide a mechanism for developing the schema of (11), e.g. we know that $\mathcal{G}/\mathrm{CUT}$ is '*at least* as powerful' as TPI, but is it 'more powerful' or equivalent (in the sense of [12])? One way in which this question could be studied is to construct translations from instances of TPI-disputes over an argument x in $\mathcal{H}$ to instances of $\varphi_{\langle\mathcal{H},x\rangle}$ of CNF-SAT. If, for φ unsatisfiable, it is then possible to simulate a derivation of $\emptyset \Rightarrow \neg\varphi$ in $\mathcal{G}/\mathrm{CUT}$ through a TPI-dispute M establishing a successful rebuttal of x in $\mathcal{H}$ with $|M|$ being comparable to the derivation length, then we could deduce that $\mathcal{G}/\mathrm{CUT}$ and TPI are equivalent. Alternatively, one might try to find example instances $\langle\mathcal{H},x\rangle$ whose dispute complexity was exponential (in $|\mathcal{X}|$) but with the proof complexity (in $\mathcal{G}/\mathrm{CUT}$) of $\varphi_{\langle\mathcal{H},x\rangle}$ being polynomial. To these ends the following translation may be of interest.

Given $\langle\mathcal{H}(\mathcal{X},\mathcal{A}),x\rangle$, the CNF-formula $\varphi_{\langle\mathcal{H},x\rangle}$ over variables $\mathcal{X}$ is,

$$x \wedge \bigwedge_{\langle y,z\rangle\in\mathcal{A}} (\neg y \vee \neg z) \;\wedge\; \bigwedge_{y\in\mathcal{X}} \left(y \;\vee \bigvee_{\{z\,:\,\langle z,y\rangle\in\mathcal{A}\}} z \right) \qquad (12)$$

It is easy to show that there is a *stable* extension of $\mathcal{H}$ containing x if and only if $\varphi_{\langle \mathcal{H},x\rangle}(\mathcal{X})$ is satisfiable.[7] The translation described in (12) suggests the following,

Problem 1. Is there a constant, k, such that for any instance $\langle \mathcal{H},x\rangle$ of a TPI-dispute wherein x is not credulously accepted,

$$\delta(\mathcal{H},x) = O\left(\pi(\emptyset \Rightarrow \{\neg\varphi_{\langle \mathcal{H},x\rangle}\}, \mathcal{G}/\text{CUT})^k\right) ?$$

Translations such as (12) also allow us to give a more precise interpretation of what might be meant by 'more powerful' dispute protocol. Thus, let Π be a (2-player) dispute protocol for argument systems (i.e. prescribing the repertoire of moves, state changes, move applicability, termination conditions, etc) with the properties that: given an instance $\langle \mathcal{H},x\rangle$ of CA

a) Π can produce a successful defence of x if and only if x is credulously accepted in $\mathcal{H}$.

b) Π either always produces a successful defence *or* always results in a successful rebuttal of x.

We can define analogous notions of *dispute complexity* with respect to arbitrary protocols – say, $\delta(\langle \mathcal{H},x\rangle,\Pi)$ – and hence regard protocol Π_1 as 'at least as powerful' as protocol Π_2 (denoted $\Pi_1 \geq \Pi_2$) if there is a constant k with which: for all dispute instances $\langle \mathcal{H},x\rangle$

$$\delta(\langle \mathcal{H},x\rangle,\Pi_1) = O(\delta(\langle \mathcal{H},x\rangle,\Pi_2)^k)$$

Problem 2. What features must be incorporated in a dispute protocol, Π, in order for it to be *more* powerful than TPI? That is, for the dispute complexity of infinitely many TPI-disputes to be superpolynomial in the dispute complexity of Π on the same instances.

Problem 3. Similarly, what features must be incorporated in Π for it to be at least as powerful as General Resolution, Gentzen Systems, etc?

It should be noted that there are subtle differences between Problems 2 and 3. The former could be examined directly without recourse to phrasing in terms of propositional proofs, the latter however is specifically concerned with the use of dispute protocols as a propositional proof mechanism.

[7] Although it is possible to construct a ('short') CNF encoding 'preferred extension containing x' rather than *stable*, this has a rather more opaque form. In any event since the absence of a preferred extension of x implies the absence of a stable extension of x, for the constructions of interest (i.e. negative instances) the TPI-dispute protocol defined still applies. Furthermore, [16, Corollary 4(a)] shows that deciding if $\mathcal{H}$ has a *stable* extension containing a given argument x is also NP-complete.

With respect to Problem 2 it has been observed earlier that something other than 'local' modifications to the state following a RETRACT move is needed.

A rather more general concern is that of what criteria must a 'reasonable' dispute protocol satisfy. From complexity-theoretic considerations, the move repertoire and its implementation cannot be permitted to be 'too powerful', e.g. treating as single operations moves which are predicated on identifying structures in an argument graph whose construction is NP-hard. While the TPI-dispute protocol is 'realistic' in the sense that the applicability of a proposed move can be validated efficiently (this, of course, is *not* the same as identifying a 'best' move), in addressing the issues raised by Problem 2 one may wish to restrict consideration to 'reasonable' protocols.[8] It is, of course, unlikely (given the conjecture NP $\neq$ CO-NP) that there is a 'reasonable' dispute protocol capable of resolving any dispute within a number of moves polynomial in the size of the argument system concerned. Nevertheless, just as the fact that existing lower bounds on Proof Complexity in failing to encompass *all possible* systems – as would be needed to prove NP $\neq$ CO-NP – motivates consideration of more powerful proof systems, so it is reasonable to examine and precisely formulate 'increasingly powerful' dispute protocols.

Finally, even for 'weak' systems such as TPI-disputes in the case of instances which lead to successful rebuttals of an argument, there is the issue of the Challenger *constructing* the 'best' line of attack, i.e. of finding the dispute that minimises dispute complexity. An analogous situation in Proof Complexity was formulated in Bonet *et al.* [9]: suppose φ is an unsatisfiable CNF with m clauses and n variables. Letting $\pi(\varphi, S)$ denote the size of the shortest proof of $\neg\varphi$ in some Propositional Proof System S, then for a function, $q : N^3 \to N$, S is said to be *q-automatizable* if there exists a (deterministic) algorithm that produces a proof (in the system S) of $\neg\varphi$ in time $q(\pi(\varphi, S), n, m)$. The cases of interest are where q is polynomially bounded in $\pi(\varphi, S)$. Informally, if a proof system is *polynomially-bounded* automatizable, then this gives an algorithm that can 'efficiently' construct a proof that is 'not much larger' than the optimal proof. The concept of q-automatizability can be reformulated in the obvious way to refer to dispute complexity (or indeed verification calculi for other CO-NP-complete problems). This motivates,

Problem 4. Let $\langle \mathcal{H}, x \rangle$ be any TPI-dispute instance in which there are n arguments and for which x is not credulously accepted in $\mathcal{H}$. Is there a deterministic algorithm that in $q(\delta(\mathcal{H}, x), n)$ steps returns a terminated TPI-dispute M establishing a successful rebuttal of x *and* with q bounded by a polynomial in $\delta(\mathcal{H}, x)$? In other words, is the TPI-dispute protocol q-automatizable for some polynomially-bounded q?

[8] Similar considerations arise in Proof Complexity and an accepted formalism has evolved to distinguish 'reasonable' from 'unreasonable' proof calculi. For the complexity-theoretic aspects affecting dispute protocols such a formalism seems a plausible basis.

To conclude our discussion of possible directions for further research, we note that our model of dispute assumes both protagonists have complete knowledge of the argument system (i.e. the finite directed graph structure). Thus the Defender may choose counterattacks which are known to eliminate particular (subsequent) attacks by the Challenger; similarly, as may be evinced by the development of the disputes from unsatisfiable CNF-formulae, the Challenger may invoke attacks, potential defences to which have been ruled out, e.g. when the Defender uses a literal y to attack a clause C, the Challenger may continue using an available clause containing $\neg y$, knowing that $\neg y$ cannot be used as a defence. In many situations it may not be the case that such complete knowledge is held *ab initio*. The modelling of disputes where the protagonists' views of the system evolve over several rounds would provide a significant development of the preliminary formalism described in this paper. Such an extension would have considerable practical interest, since many of the implementations require such evolution. For example, Gordon's [19] game is intended to induce the particpants to present the arguments that they wish to deploy, essentially establishing the argumentation framework which will subsequently used when the question comes to trial. In [6] it is assumed that each participant has only a partial view on the argumentation framework which is extended by elements recognised by their opponents as the dialogue proceeds. If we consider disputes between autonomous agents, it is perhaps unrealistic to expect them to begin with a shared understanding of the overall argumentation framework.

5. Conclusion

In this paper we have introduced a formal concept of dispute complexity with which to consider questions regarding the number of rounds a 2-player dialogue concerning a particular argument, may require before one player accepts that the argument is/is not defensible. Building on the Argument System formalism of [15] and the argument game - TPI-dispute - discussed in [34], a precise formulation of the latter has been presented. With this formulation at hand, we are able to prove that there are instances representing a win for the Challenger but for which exponentially many rounds must be played before the Defender is convinced of this. Our techniques exploit the close relationship between such dispute protocols and the concept of formal proof calculi for propositional tautologies by showing that the TPI-dispute protocol applied to representations of these can be used to build a proof of validity in a CUT-free Gentzen System whose length is comparable to the number of rounds needed in a TPI-dispute. The ideas and techniques put forward in this paper represent just a preliminary foundation: an extensive range of open questions and further directions for research arise from this, only a selection of which have been discussed in detail.

References

1. Ajtai, M.: 1988, 'The Complexity of the Pigeonhole Principle'. In: *29th Annual Symposium on Foundations of Computer Science*. White Plains, New York, pp. 346–355.
2. Avron, A.: 1993, 'Gentzen-type Systems, Resolutions, and Tableaux'. *Journal of Automated Reasoning* **10**(2), 265–281.
3. Beame, P., R. Impagliazzo, J. Krajíček, T. Pitassi, P. Pudlák, and A. Woods: 1992, 'Exponential Lower Bounds for the Pigeonhole Principle'. In: *Proceedings of the Twenty-Fourth Annual ACM Symposium on the Theory of Computing*. Victoria, British Columbia, Canada, pp. 200–220.
4. Beame, P. and T. Pitassi: 1996, 'An exponential separation between the parity principle and the pigeonhole principle'. *Annals of Pure and Applied Logic* **80**(3), 195–228.
5. Beame, P. and T. Pitassi: 1998, 'Propositional Proof Complexity: Past, Present, and Future'. *Bulletin of the EATCS* **65**, 66–89.
6. Bench-Capon, T. J. M.: 1998, 'Specification and Implementation of Toulmin Dialogue Game'. In: J. C. H. *et al.* (ed.): *Legal Knowledge Based Systems*. pp. 5–20.
7. Berge, C.: 1973, *Graphs and Hypergraphs*. Amsterdam: North-Holland.
8. Bondarenko, A., P. M. Dung, R. A. Kowalski, and F. Toni: 1997, 'An Abstract, Argumentation-Theoretic Approach to Default Reasoning'. *Artificial Intelligence* **93**(1–2), 63–101.
9. Bonet, M., T. Pitassi, and R. Raz: 1997, 'Lower bounds for cutting planes proofs with small coefficients'. *Journal of Symbolic Logic* **62**(3), 708–728.
10. Brewka, G.: 1994, 'A Reconstruction of Rescher's Theory of Formal Disputation Based on Default Logic'. In: A. G. Cohn (ed.): *Proceedings of the Eleventh European Conference on Artificial Intelligence*. Chichester, pp. 366–370.
11. Chvátal, V.: 1977, 'Determining the Stability Number of a Graph'. *SIAM Journal on Computing* **6**(4), 643–662.
12. Cook, S. A. and R. A. Reckhow: 1979, 'The Relative Complexity of Propositional Proof Systems'. *Journal of Symbolic Logic* **44**(1), 36–50.
13. Dimopoulos, Y., B. Nebel, and F. Toni: 1999, 'Preferred Arguments are Harder to Compute than Stable Extensions'. In: D. Thomas (ed.): *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI-99-Vol1)*. San Francisco, pp. 36–43.
14. Dimopoulos, Y., B. Nebel, and F. Toni: 2000, 'Finding Admissible and Preferred Arguments Can Be Very Hard'. In: A. G. Cohn, F. Giunchiglia, and B. Selman (eds.): *KR2000: Principles of Knowledge Representation and Reasoning*. San Francisco, pp. 53–61.
15. Dung, P. M.: 1995, 'On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reason, Logic Programming, and N-Person Games'. *Artificial Intelligence* **77**, 321–357.
16. Dunne, P. E. and T. J. M. Bench-Capon: 2001, 'Complexity and Combinatorial Properties of Argument Systems'. Technical report, Dept. of Comp. Sci., Univ. of Liverpool.
17. Fisher, M.: 1991, 'A Resolution Method for Temporal Logic'. In: R. Myopoulos and J. Reiter (eds.): *Proceedings of the 12th International Joint Conference on Artificial Intelligence*. Sydney, Australia, pp. 99–104.

18. Gentzen, G.: 1969, ‘Investigations into Logical Deductions, 1935’. In: M. E. Szabo (ed.): *The Collected Papers of Gerhard Gentzen*. Amsterdam: North-Holland Publishing Co., pp. 68–131.
19. Gordon, T. F.: 1995, *The Pleadings Game: An Artificial Intelligence Model of Procedural Justice*. Dordrecht: Kluwer Academic Publishers.
20. Haken, A.: 1985, ‘The intractability of resolution’. *Theoretical Computer Science* **39**(2-3), 297–308.
21. Jakobovits, H. and D. Vermeir: 1999, ‘Dialectic Semantics for Argumentation Frameworks’. In: *Proceedings of the Seventh International Conference on Artificial Intelligence and Law (ICAIL-99)*. N.Y., pp. 53–62.
22. Lodder, A. R.: 1998, ‘Dialaw: On legal justification and Dialogue Games’. Ph.D. thesis, Univ.of Maastricht.
23. Loui, R. P.: 1998, ‘Process and Policy: Resource-Bounded Nondemonstrative Reasoning’. *COMPINT: Computational Intelligence: An International Journal* **14**, 1–38.
24. MacKenzie, J. D.: 1978, ‘Question-Begging in Non-Cumulative Systems’. *Journal of Philosophical Logic* **8**(1), 117–133.
25. Pitassi, T. and A. Urquhart: 1995, ‘The complexity of the Hajós calculus’. *SIAM Journal on Discrete Mathematics* **8**(3), 464–483.
26. Prakken, H. and G. Sartor: 1997, ‘Argument-Based Extended Logic Programming with Defeasible Priorities’. *Journal of Applied Non-classical Logics* **7**, 25–75.
27. Pudlák, P.: 1998, ‘The Lengths of Proofs’. In: S. R. Buss (ed.): *Handbook of Proof Theory*, Vol. 137 of *Studies in Logic and the Foundations of Mathematics*. Amsterdam: North-Holland, Chapt. 8, pp. 547–637.
28. Pudlák, P.: 1999, ‘On the complexity of propositional calculus’. In: *Sets and Proofs, Invited papers from Logic Colloquium’97*. pp. 197–218.
29. Reed, C.: 1998, ‘Dialogue frames in agent communications’. In: Y. Demazeau (ed.): *Proc. 3rd International Conference on Multi-agent systems (ICMAS-98)*. pp. 246–253.
30. Toulmin, S.: 1959, *The uses of argument*. Cambridge University Press.
31. Urquhart, A.: 1989, ‘The complexity of Gentzen systems for propositional logic’. *Theoretical Computer Science* **66**(1), 87–97.
32. Urquhart, A.: 1995, ‘The complexity of propositional proofs’. *Bulletin of Symbolic Logic* **1**(4), 425–467.
33. Vreeswijk, G.: 1993, ‘Defeasible Dialectics: A Controversy-Oriented Approach towards Defeasible Argumentation’. *Journal of Logic and Computation* **3**(3), 317–334.
34. Vreeswijk, G. and H. Prakken: 2000, ‘Credulous and sceptical argument games for preferred semantics.’. In: *Proceedings of JELIA’2000, The 7th European Workshop on Logic for Artificial Intelligence.* Berlin, pp. 224–238.
35. Walton, D. N. and E. C. W. Krabbe: 1995, *Committment in Dialogue: Basic Concepts of Interpersonal Reasoning*. Univ. of New York Press.