Argumentation

Based on the material due to

P. M. Dung, R.A. Kowalski et al.

1. Introduction
Argumentation constitutes a major component of human's intelligence. The ability to engage in arguments is essential for humans to understand new problems, to perform scientific reasoning, to express, clarify and defend their opinions in their daily lives.

The way humans argue is based on a very simple principle which is summarized succinctly by an old saying

"*The one who has the last word laughs best*".

To illustrate this principle, let us take a look at an example, a mock argument between two persons I and A, whose countries are at war, about who is responsible for blocking negotiation in their region.

Example 1 Consider the following dialog

I: My government can not negotiate with your government because your government doesn’t even recognize my government

A: Your government doesn’t recognize my government either

Consider the following continuation of the above arguments:

I: But your government is a terrorist government

This represents an attack on the A’s argument. If the exchange stops here, then I clearly has the "last word", which means that he has successfully argued that A’s government is responsible for blocking the negotiation.

Part 1: Acceptability of Arguments

Definition 1 An *argumentation framework* is a pair

$$AF = \langle AR, attacks \rangle$$

where AR is a set of arguments, and attacks is a binary relation on AR, i.e. $attacks \subseteq AR \times AR$.

For two arguments A,B, the meaning of attacks(A,B) is that A represents an attack against B.

The dialog between persons I and A from Example 1 can be represented by an argumentation framework <AR,attacks> as follows:

$$AR = \{i1,i2,a\} \text{ and } attacks = \{(i1,a),(a,i1),(i2,a)\}$$

with i1,i2 denoting the first and the second argument of I, respectively, and a denoting the argument of A.

We can depict the end of the above dialog by the following oriented graph

$$\boxed{i1 \leftarrow a \leftarrow i2}$$

Fig. 1

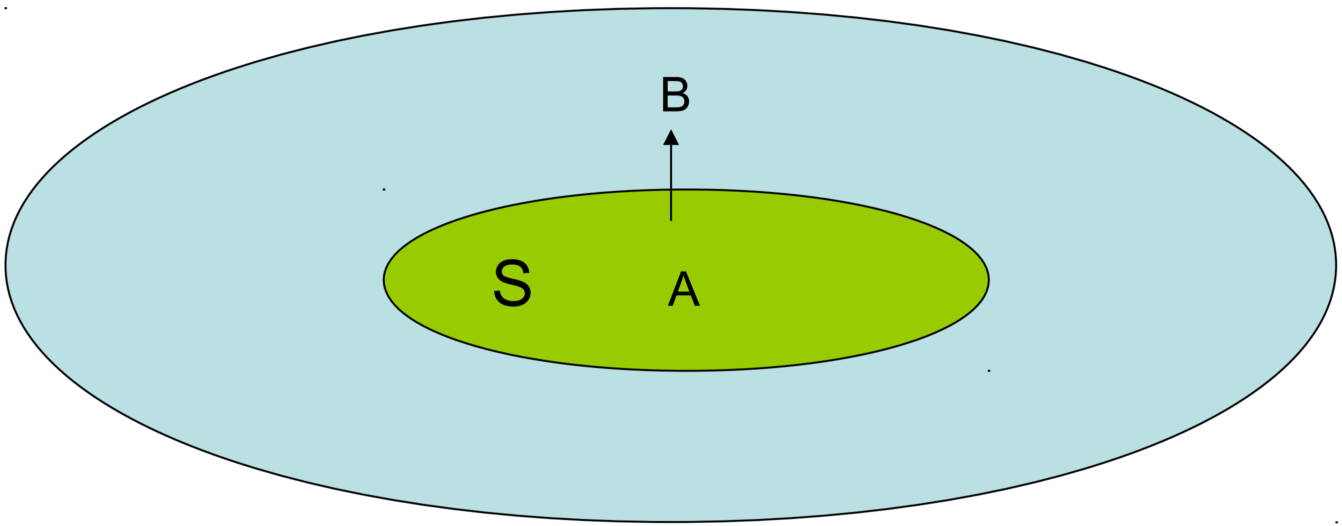
Remark 1 From now on, if not explicitly mentioned otherwise, we always refer to an arbitrary but fixed argumentation framework

$$AF = \langle AR, attacks \rangle.$$

(i) Further, we say that A attacks B (or B is attacked by A) if attacks(A,B) holds.

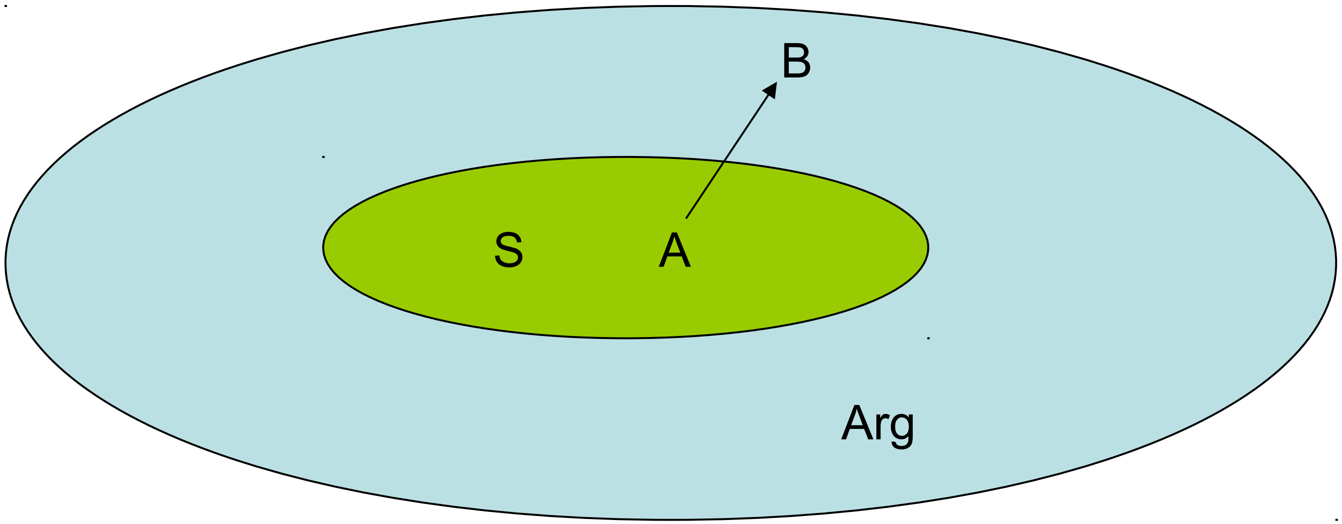
(ii) Similarly, we say that a set S of arguments attacks B $\in$ Arg (or B is attacked by S) if B is attacked by an argument in S i.e.

$$S \text{ attacs } B \Leftrightarrow (\exists A \in S)(attacs(A,B))$$



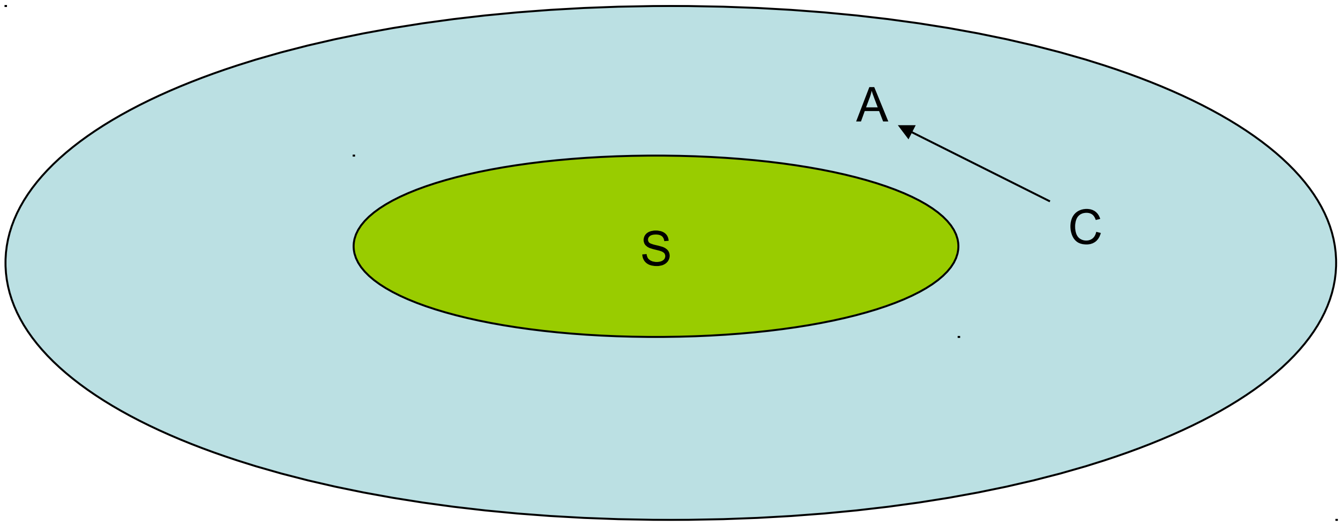
Definition 2 (conflict-free sets, acceptable arguments)

(i) A set S of arguments is said to be *conflict-free* if there are no arguments A,B in S such that A attacks B .

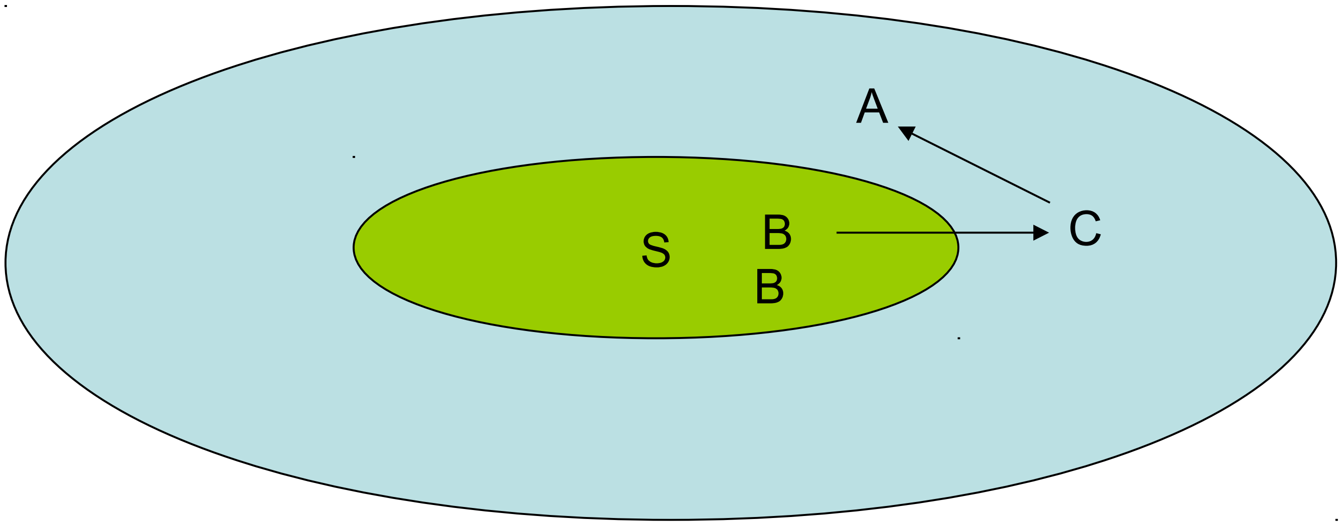


$$(A \in S \quad \& \quad attack(A,B)) \quad => B \notin A$$

(ii) we say that an argument A is acceptable with respect to a set S if S can defend A by an argument $B \in S$ against all attacks C on A.



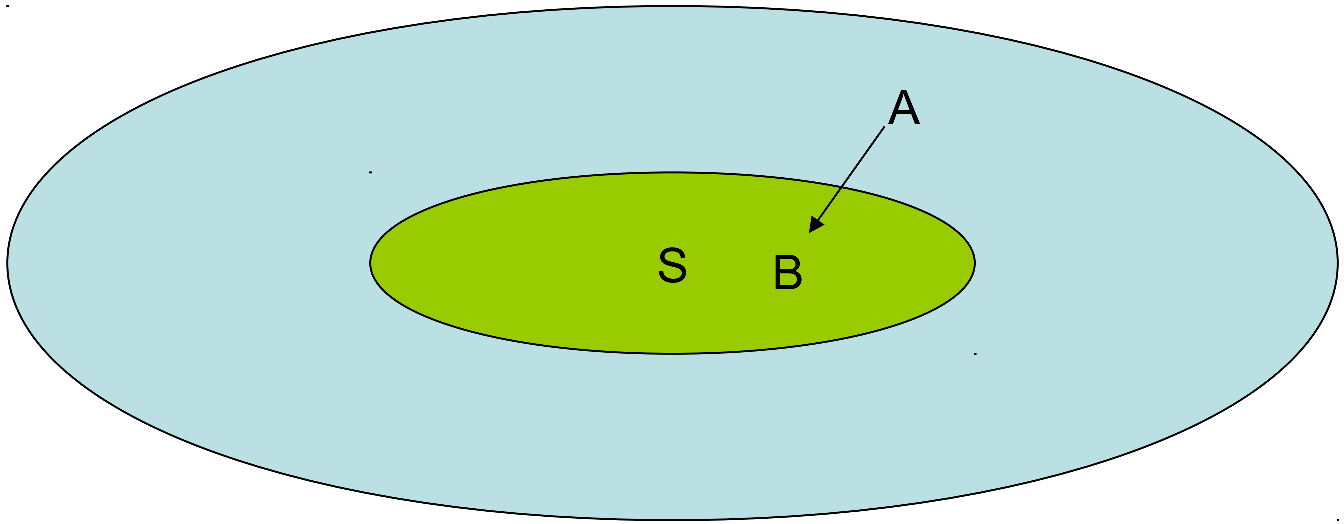
$(A \in Arg \ \& \ attacks(C,A))$



$(A \in Arg\ \ \&\ \ attacks(C,A)) \Rightarrow (\exists B \in S)\ (attacks(B,C))$

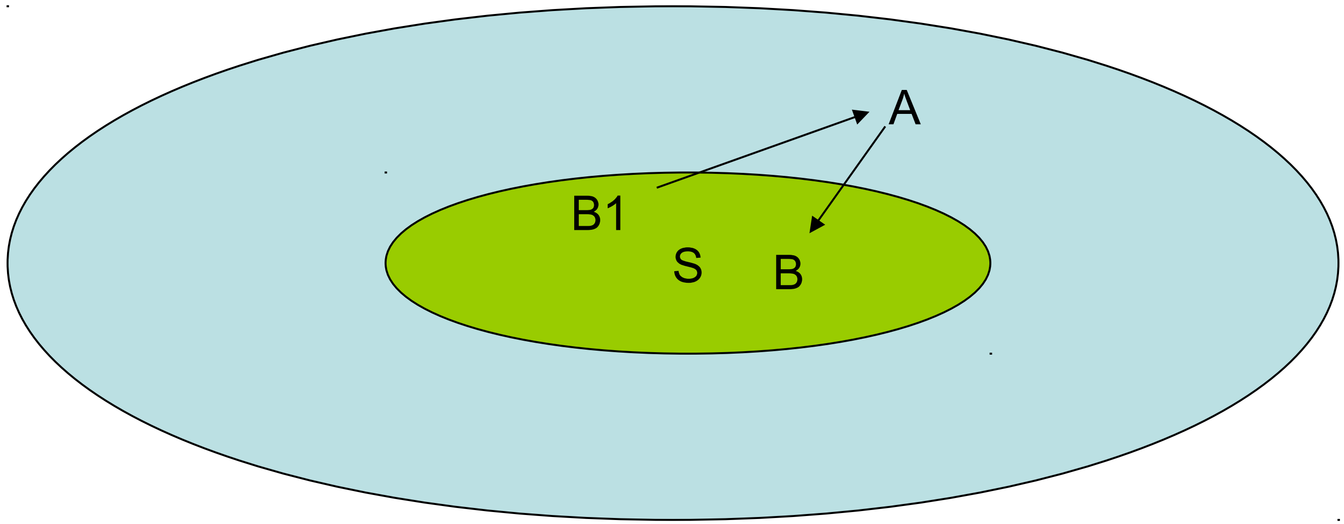
Definition 3 (admissible of sets of arguments)

(i) A conflict-free set of arguments S is *admissible* iff each argument in S is acceptable wrt S i.e. S can defend all its arguments.



$B \in S$ & $A \in Arg$ & attacs(A,B)

Note that a conflict-free set of arguments S is *admissible* iff each argument in S is acceptable wrt. S i.e. S can defend all its arguments.
Admissible sets are good prototypes for sets of arguments given to reasoning agents.



$(B \in S \ \& \ A \in Arg \ \& \ attacs(A,B)) \Rightarrow (\exists B1 \in S)(attacs(B1,A))$

The (credulous) semantics for Argumentation

of an argumentation framework is defined by the notion of preferred extesions.

Definition 4 A *preferred extension* of an argumentation framework AF is a maximal (wrt. set inclusion) admissible set of AF.

(Continuation of Example 1). It is not difficult to see that the correspon-ding AF has exactly one prefered extension

$$\{i1,i2\} \subseteq \text{Arg}$$

Indeed, as depicted in Fig. 1

$$i1 \leftarrow a \leftarrow i2$$

the only argument attacking i1 is a , but it is attacked by i2.
And i2 no argument I attacking i2.

Example 2 (Nixon Diamond). The well-known Nixon diamond example can be represented as an argumentation framework AF = <AR,attacks> with

AR = {A,B}, and attacks = {(A,B),(B,A)}

where A represents the argument "Nixon is anti-pacifist since he is a republican", and B represents the argument "Nixon is a pacifist since he is a quaker".

This argumentation framework has two preferred extensions

{A} , {B}

one in which Nixon is a pacifist and one in which Nixon is a quaker.

It is not difficult to see that if an argumentation framework consist a circle,

e.g. Attacs = {(A,B),(B,C),(C,A),} then it has the preferred extension consisting of

{A} , {B} , {C}

Lemma 1 (Fundamental Lemma)

Let S be an admissible set of arguments and A , A' be arguments which are acceptable wrt S. Then

(i) $S' = S \cup \{A\}$ is admissible, and

(ii) A' is acceptable wrt S'.

Proof (i) We need only to show that S' is conflict-free. Assume the contrary. Therefore, there exists an argument $B \in S$ s.t.

A attacks B or B attacks A.

From the admissibility of S and the acceptability of A, there is an argument $B' \in S$ such that

B' attacks B or B' attacks A.

Since S is conflict-free, it follows that B' attacks A. But then there is an argument B" in S s.t.

B" attacks B'. Contradiction !!

(ii) Obvious.

The following theorem follows directly from the fundamental lemma.

Theorem 1 Let AF be an argumentation framework.

(i) The set of all admissible sets of AF form a complete partial order wrt. set inclusion.

(ii) For each admissible set S of AF, there exists an preferred extension E of AF such that $S \subseteq E$.

Theorem 1 together with the fact that the empty set is always admissible, implies the following corollary:

Corollary 2 Every argumentation framework possesses at least one preferred extension.

Hence, preferred extension semantics is always defined for any argumentation framework.

Stable Semantics for Argumentation

We introduce the following notion of stable extension and show that grounded stable extensions provide so called skeptical semantics.

Definition 5 A conflict-free set of arguments S is called *stable extension* iff S attacks each argument which does not belong to S.

Formally,

$$(\forall A \in Arg)\ [A \notin S \Rightarrow (\exists B \in S)(attacks(B,A)]$$

We will show later on that in the context of game theory, our notion of stable extension coincides with the notion of stable solutions of n-person games introduced by Von Neuman and Morgenstern in 1944.

It is easy to see that the following holds

Lemma 3 S is a stable extension iff S = { A | A is not attacked by S}.

It will turns out later that this proposition underlines exactly the way the notions of

- stable models in logic programming,
- extensions in Reiter’s default logic, and
- stable expansion in Moore’s autoepistemic logic are defined.

The relations between stable extension and preferred extension are clarified in the following lemma.

Lemma 4 Every stable extension is a preferred extension, but not vice versa.

Proof It is clear that each stable extension is a preferred extension. To show that the reverse does not hold, we construct the following argumentation framework:

AF = (AR,attacks) with AR = {A} and attacks = {(A,A)}.

Here, the empty set is a preferred extension of AF which is clearly not stable.

On the other hand, in examples 1 and 4, preferred extensions and stable extension semantics coincide.

Though stable semantics is not defined for every argumentation system, an often asked question is whether or not argumentation systems with no stable extensions represent meaningful systems ???

In part (2) of this paper, we will provide meaningful argumentation systems without stable semantics, and thus provide an definite answer to this question.

Fixpoint Semantics and Grounded (Sceptical) Semantics
We show in this chapter that argumentation can be characterized by a fixpoint theory providing an elegant way to introduce grounded (skeptical) semantics.

Definition 6 The *characteristic function* of an argumentation framework AF, denoted by F_{AF}, = <AR,attacks> is defined as follows:

$$F_{AF}: 2^{AR} \to 2^{AR}$$
$$F_{AF}(S) = \{ A \mid A \text{ is acceptable wrt } S \}$$

Notation As we always refer to an arbitrary but fixed argumentation framework AF, we often write F instead of F_{AF} for short.

Lemma 5 A conflict-free set S of arguments is admissible iff $S \subseteq F(S)$.

Proof The lemma follows immediately from the property

"If S is conflict-free then F(S) is also conflict-free".

So we need only to prove this property. Assume that there are A,A' in F(S) such that A attacks A'. Thus, there exists B is S such that B attacks A. Hence there is B' in S such that B' attacks B. Contradiction !!
So F(S) is conflict free.

It is easy to see that if an argument A is acceptable wrt S then A is also acceptable wrt. any superset S' of S. From this fact follows immediately the next lemma.

Lemma 6 F_{AF} is monotonic (wrt. set inclusion).

Definition 7 An argumentation framework AF = <AR,attacks> is *finitary* iff for each argument A, there are only finitely many arguments in AR which attack A.

Lemma 7 If AF is finitary then F_{AF} is ω-continuous.

Proof Let $S_0 \subseteq \ldots \subseteq S_n \subseteq \ldots$ be an increasing sequence of sets of arguments, let

$$S = \cup\{ S_i \mid i \in N\}$$

Let $A \in F_{AF}(S)$. Since there are only finitely many arguments which attack A, there exists a number m such that $A \in F_{AF}(S_m)$.

Therefore,

$$F_{AF}(S) = \cup\{F_{AF}(S_i) \mid i \in N\}$$

and F_{AF} is ω-continuous.

Consequently, it follows from the Theorem of Knaster and Tarski that function F_{AF} has the least fix point.

Thus the following definition is justified.

Definition 8 The *grounded extension* of an argumentation framework AF, denoted by GE_{AF}, is the least fixed point of F_{AF}.

Example 3 (Continuation of example 1)

It is easy to see: $F_{AF}(\varnothing) = \{i2\}$,

$F_{AF}{}^{2}(\varnothing) = \{i1,i2\}$,

$F_{AF}{}^{3}(\varnothing) = F_{AF}{}^{2}(\varnothing)$

…

Thus $GE_{AF} = \{i1,i2\}$. Note that GE_{AF} is also the only preferred extension of AF.

Example 4 (Continuation of the Nixon-example).
Let

AF = <AR,attacks> with AR = {A,B} and attacks = {(A,B),(B,A)}

Then it follows immediately that the grounded extension is empty, i.e. a sceptical reasoner will not conclude anything.

The following notion of complete extension provides the link between preferred extensions (credulous semantics), and grounded extension (sceptical semantics).

Definition 9 An admissible set S of arguments is called a *complete extension* iff each argument which is acceptable wrt. S, belongs to S.

Intuitively, the notion of complete extensions captures the kind of confident rational agents who believes in everything he can defend.

Lemma 8 A conflict-free set of arguments E is a complete extension iff $E = F_{AF}(E)$.

Proof Note that

$$F_{AF}(E) = \{A \mid A \text{ is acceptable wrt. } E\}$$

The relations between preferred extensions, grounded extensions and complete extensions is given in the following theorem.

Definition 8 The *grounded extension* of an argumentation framework AF, denoted by GE_{AF}, is the least fixed point of F_{AF}.

Example 5 (Continuation of example 3). It is easy to see:

$$F_{AF}(\varnothing) = \{i2\}$$
$$F_{AF}^{2}(\varnothing) = \{i1,i2\}$$
$$F_{AF}^{3}(\varnothing) = \{i1,i2\} = F_{AF}^{2}(\varnothing)$$

Thus $GE_{AF} = \{i1,i2\}$. Note that GE_{AF} is also the only preferred extension of AF.

The relations between preferred extensions, grounded extensions and complete extensions is given in the following theorem.

Theorem 2 (i) Each preferred extension is a complete extension, but not vice versa.
(ii) The grounded extension is the least (wrt set inclusion) complete extension.
(iii) The complete extensions form a complete semilattice wrt. set inclusion.

Proof (i) It is obvious from the fixpoint definition of complete extensions that every preferred extension is a complete extension.

The Nixon diamond example provides a counter example that the reverse does not hold since the empty set is a complete extension but not a preferred one.

(ii) Obvious
(iii) Let SE be a non-empty set of complete extensions.
Let LB = { E | E is admissible and E $\subseteq$ E' for each E' in SE }.

It is clear that GE $\in$ LB. So LB is not empty. Let S = E{ E | E $\in$ LB}.
It is clear that S is admissible, i.e. S $\subseteq$ F(S). Let E = lub(F^i(S)) for ordinals i.
Then it is clear that E is a complete extension and E $\in$ LB.
Thus E = S. So E is the glb of SE.

Remark In general, the intersection of all preferred extensions does not coincide with the grounded extension.

Sufficient Conditions for Coincidence between Different Semantics

Well-Founded Argumentation Frameworks

We want to give in this paragraph a sufficient condition for the coincidence between the grounded semantics and preferred extension semantics as well as stable semantics.

Definition 9 An argumentation framework is *well-founded* iff there exists no infinite sequence $A_0, A_1, \ldots, A_n, \ldots$ such that for each i, A_{i+1} attacks A_i.

The following theorem shows that well-founded argumentation frameworks have exactly one extension.

Theorem 3 Every well-founded argumentation framework has exactly one complete extension which is grounded, preferred and stable.

Proof Assume the contrary, i.e. there exist a well-founded argumentation framework whose grounded extension is not a stable extension.

Let AF=(AR,attacks) be such a argumentation framework where

$$S = \{ A \mid A \in AR\text{-}GE_{AF} \text{ and } A \text{ is not attacked by } GE_{AF} \}$$

is nonempty.

Now

we want to show that each argument A in S is attacked by S itself.

Let $A \in S$. Since A is not acceptable wrt GE_{AF}, there is an attack B against A such that B is not attacked by GE_{AF}. From the definition of S, it is clear that B does not belong to GE_{AF}. Hence, B belongs to S.

Thus there exists an infinite sequence $A_1, A_2, \ldots$ such that for each **i**, A_{i+1} attacks A_i. Contradiction.

Coherent Argumentation Frameworks

Now, we want to give a condition for the coincidence between stable extensions and preferred extensions.

In general, the existence of a preferred extension which is not stable indicates the existence of some "anomalies" in the corresponding argumentation framework.

For example, the argumentation framework

$$<\{A\},\{(A,A)\}>$$

has an empty preferred extension which is not stable. Note that this argumentation framework corresponding to the logic program $p \leftarrow \text{not } p$ is of this kind.

So it is interesting to find sufficient conditions to avoid such anomalies.

Definition 10 (i) An argumentation framework AF is said to be *coherent* if each preferred extension of AF is stable.

(ii) We say that an argumentation framework AF is *relatively grounded* if its grounded extension coincides with the intersection of all preferred extensions.

It follows directly from the definition that there exists at least one stable extension in a coherent argumentation framework.

Motivation. Imagine an exchange of arguments between you and me about some proposition C. You start by putting forward an argument A_0 supporting C. I don't agree with C, and so I present an argument A_1 attacking your argument A_0.

To defend A_0 and so C, you put forward another argument A_2 attacking my argument A1. Now I present A_3 attacking A_2.

If we stop at this point, A_0 is defeated. It is clear that A_3 plays a decisive role in the defeat of A_0 though A_3 does not directly attack A_0.

It is said that A_3 represents an indirect attack against A_0.

In general, we say that an argument B *indirectly attacks* A if there exists a finite sequence $A_0,..,A_{2n+1}$ such that

(i) $A = A_0$ and $B = A_{2n+1}$, and
(ii) for each i, $0 \leq i \leq 2n$, A_{i+1} attacks A_i.

An argument B is said to be *controversial* wrt. A if B indirectly attacks A and indirectly defends A. Hence if some C attacks A We have the following picture

B →> A ← C <← B

An argument is *controversial* if it is controversial wrt. some argument A.

Definition 11 (i) An argumentation framework is *uncontroversial* if none of its arguments is controversial.
(ii) An argumentation framework is *limited controversial* if there exists no infinite sequence of arguments

$$A_0,\ldots,A_n,\ldots \text{ such that } A_{i+1} \text{ is controversial wrt } A_i.$$

It is clear that every uncontroversial argumentation framework is limited controversial but not vice versa.

Theorem 4 (i) Every limited controversial argumentation framework is coherent.
(ii) Every uncontroversial argumentation framework is coherent and relatively grounded.

Proof follows from the following lemmas 9,10.

For the proof of lemmas 9,10, we need a couple of new notations. We shall not prove these lemmas, but the concepts are intersting.

An argument A is said to be a *threat* to a set of argument S if A attacks S and A is not attacked by S.
A set of arguments D is called a *defense* of a set of argument S if D attacks each threat to S.

Lemma 9 Let AF be a limited controversial argumentation framework. Then there exists at least a nonempty complete extension E of AF.

Lemma 10 Let AF be an uncontroversial argumentation framework, and A be an argument such that A is not attacked by the grounded extension GE of AF and $A \in GE$.
Then
(i) there exists a complete extension E1 such that $A \in E1$, and
(ii) there exists a complete extension E2 such that E2 attacks A.

Corollary 11 Every limited controversial argumentation framework possesses at least one stable extension.

This corollary in fact gives the answer to an often asked question about the existence of stable semantics of knowledge representation formalisms like Reiter's default logic, logic programming or autoepistemic logic.