An Introduction to Formal Argumentation

Martin Caminada

University of Luxembourg

An Example (1/2) [Prakken]

Paul: My car is very safe.

Olga: Why?

Paul: Since it has an airbag.

Olga: It is true that your car has an airbag, but I do not think that this makes your car safe, because airbags are unreliable: the newspapers had several reports on cases where airbags did not work.

Paul: I also read that report but a recent scientific study showed that cars with airbags are safer than cars without airbags, and scientific studies are more important than newspaper reports.

Olga: OK, I admit that your argument is stronger than mine. However, your car is not very safe, since its maximum speed is much too high.

Arguments and attacks

Argument: expresses one or more reasons that lead to a proposition

$a, b, c \Rightarrow d$ or $a, b \Rightarrow c; \; c \Rightarrow d$

An argument can *attack* another argument

rebutting attack:

attack one of the *conclusions* of the other argument:

$e, f, g \Rightarrow \neg d$ against $a, b, c \Rightarrow d$

undercutting attack:

attack the *reasons* of the other argument

$e, f, g, \Rightarrow [a, b, c \not\Rightarrow d]$against $a, b, c \Rightarrow d$

Example (2/2)

A: My car is very safe, since it has an airbag:

$$\text{has\_airbag} \Rightarrow \text{safe}$$

B: The newspapers say that airbags are not reliable, so having an airbag is not a good reason why your car is safe

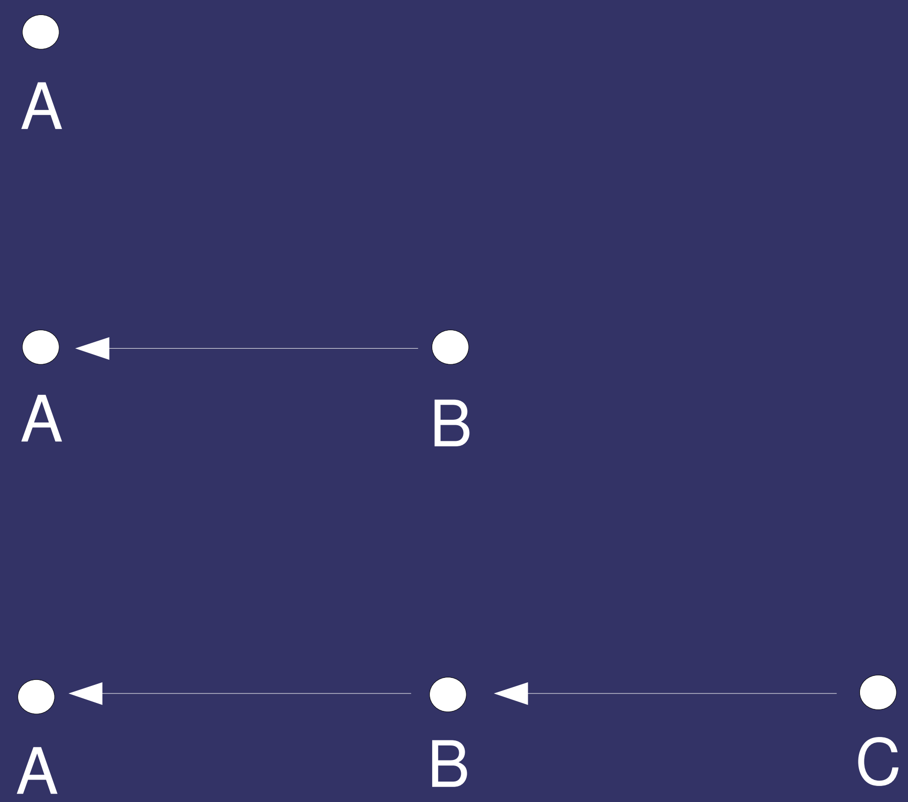
$$\text{say}(\text{npr}, \neg\text{rel}(\text{airbag})) \Rightarrow \neg\text{rel}(\text{airbag})$$

$$\neg\text{rel}(\text{airbag}) \Rightarrow [\text{has\_airbag} \not\Rightarrow \text{safe}]$$

C: Scientific reports say that airbags are reliable.

$$\text{say}(\text{sr}, \text{rel}(\text{airbag})) \Rightarrow \text{rel}(\text{airbag})$$

How Arguments Interact (1/2)

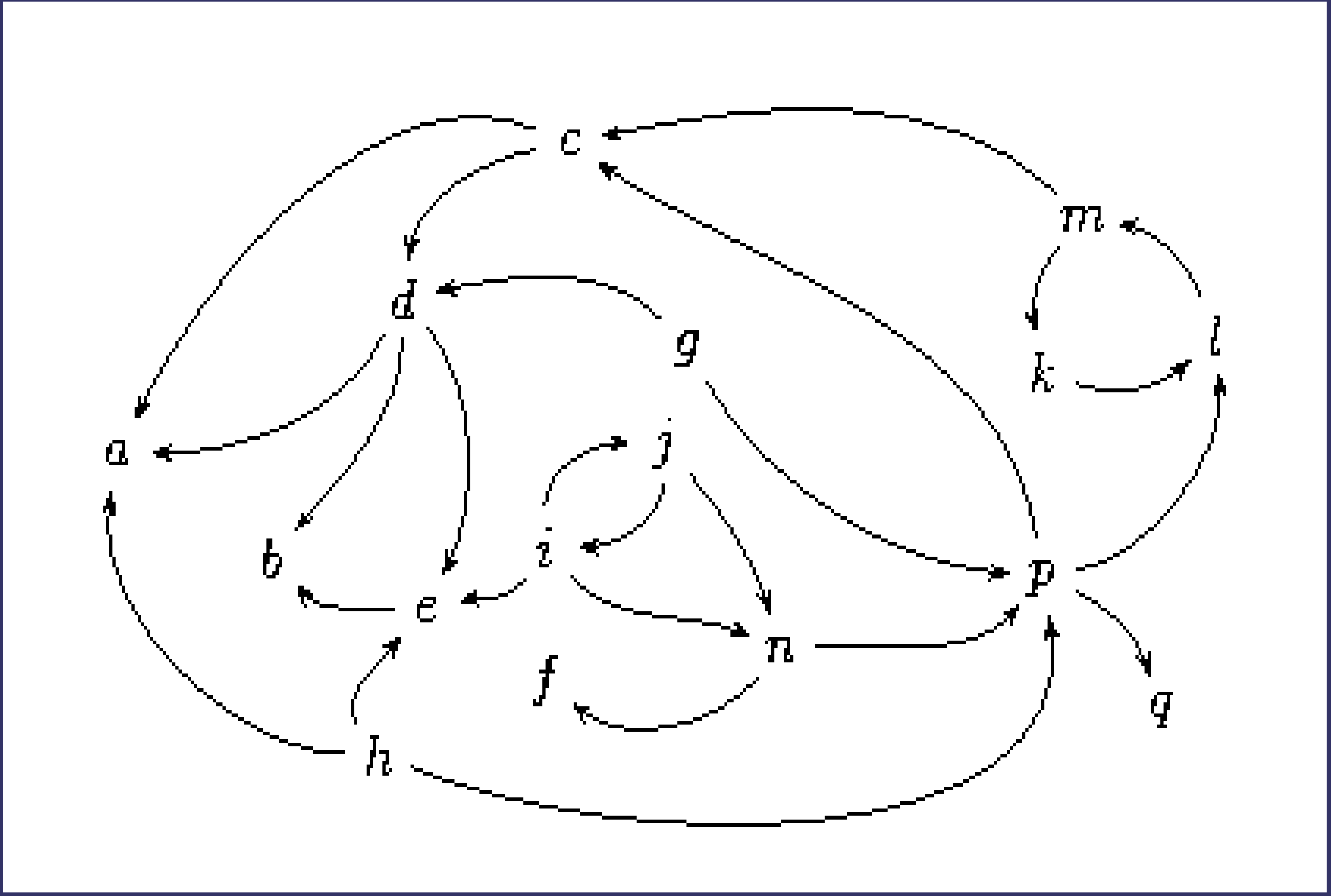


A: my car is very safe since it has an airbag

B: newspapers say that airbags are unreliable

C: scientific reports say that airbags are reliable, and these are more important than newspapers

How Arguments Interact (2/2)



Argumentation: what is it good for?

Legal reasoning: CATO/HYPO
use argumentation tools for supporting lawyers

Medical reasoning: CRUK/CARREL
helping doctors to suggest the best
treatment for their patients

Nonmonotonic Logic

$$\Phi \vdash \varphi$$

$$\not\Rightarrow$$

$$\Phi \cup \Psi \vdash \varphi$$

The Argumentation Approach

generate arguments based on a knowledge base

see how these arguments defeat each other

determine which arguments can be seen as justified

take the conclusions of the justified arguments

Argumentation in Agent Systems

For <u>internal reasoning</u> of single agents

reasoning about beliefs, goals, intentions etc is often defeasible

For <u>interaction</u> between multiple agents

information exchange involves explanation

collaboration and negotiation involve conflict of opinion and persuasion

What Arguments Look Like (1/2)
Arguments as Sets of Assumptions

Given a knowledge base (K, Ass)

Argument: (A, c) with $A \subseteq Ass$ s.t.:

$A \cup K \models c$

$A \cup K \not\models \bot$

$\nexists\, a \in A:\ A \setminus \{a\} \cup K \models c$

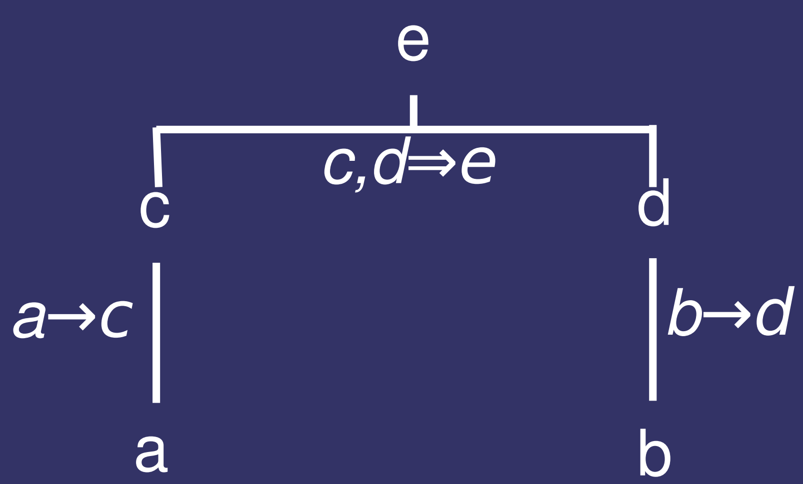
(Besnard & Hunter, 2001)

What Attacks Look Like (1/2)
Arguments as Sets of Assumptions

Assumption attack:

(A_2, c_2) attacks (A_1, c_1) iff $\neg c_2 \in A_1$

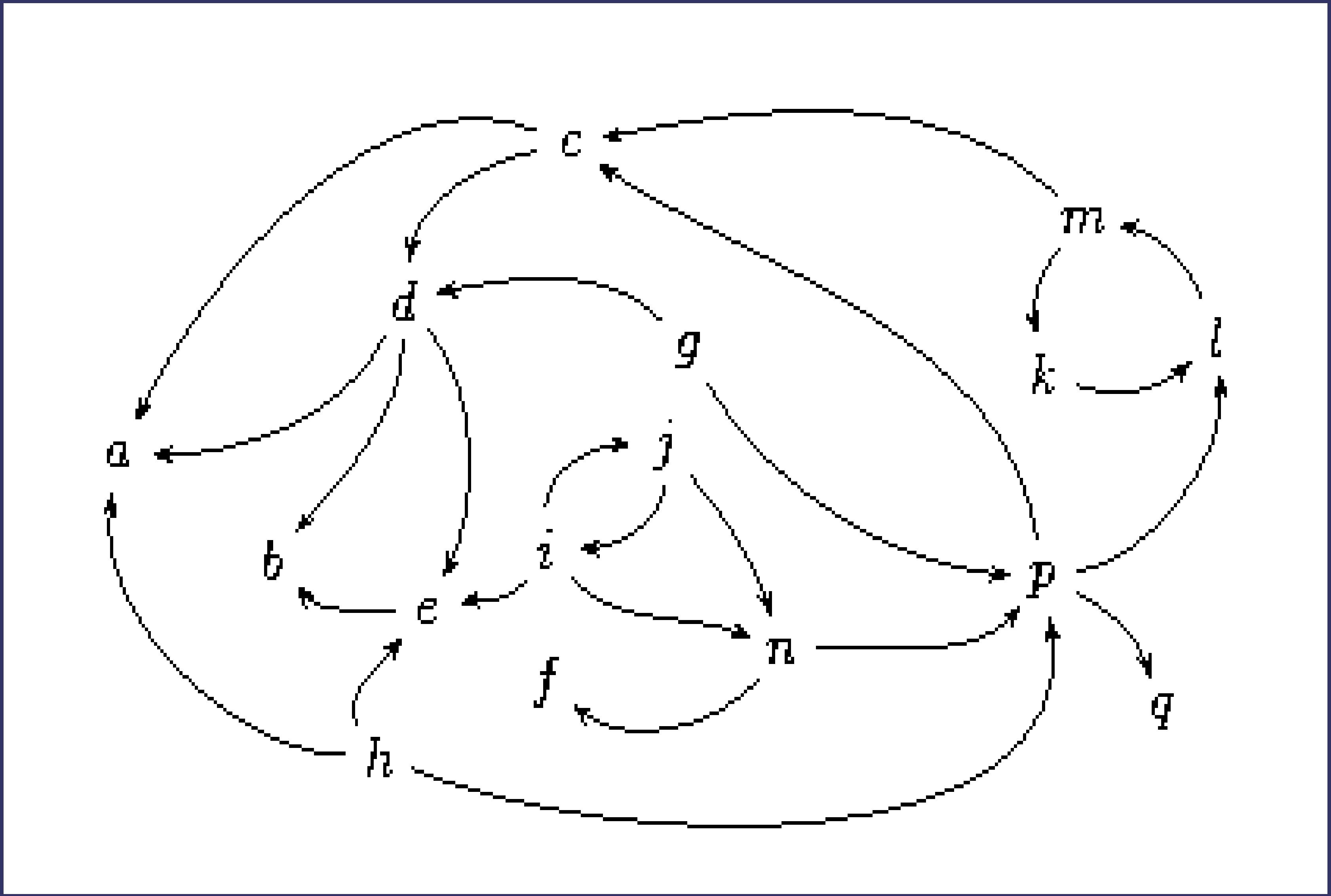
What Arguments Look Like (1/2)
Arguments as Trees Constructed with Rules



$((a) \rightarrow c), ((b) \rightarrow d) \Rightarrow e$

strict rule (→): “from ... it always follows that...”
defeasible rule (⇒): “from ... it usually follows that...”

How Arguments Interact (2/2)



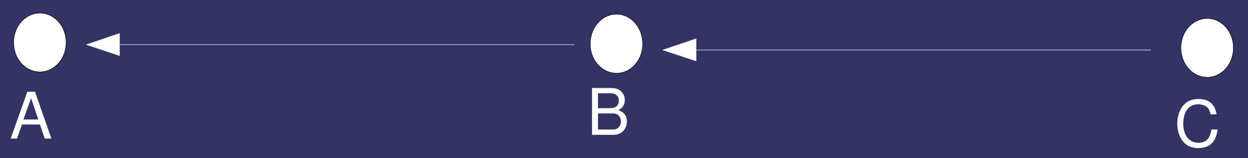
Argument Evaluation Postulate

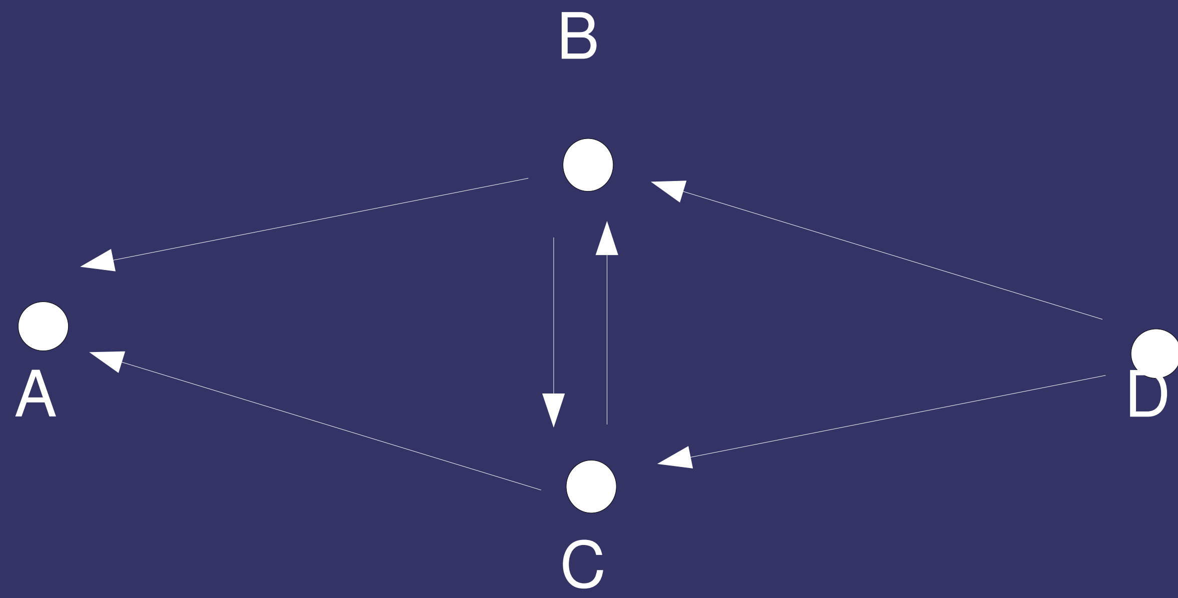
argument labels: in, out, undec

An argument is in
iff all its defeaters are out

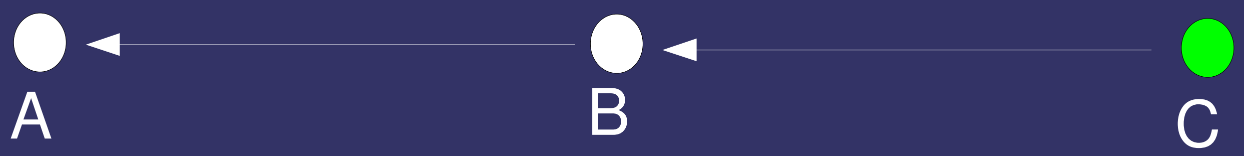
An argument is out
iff it has a defeater that is in

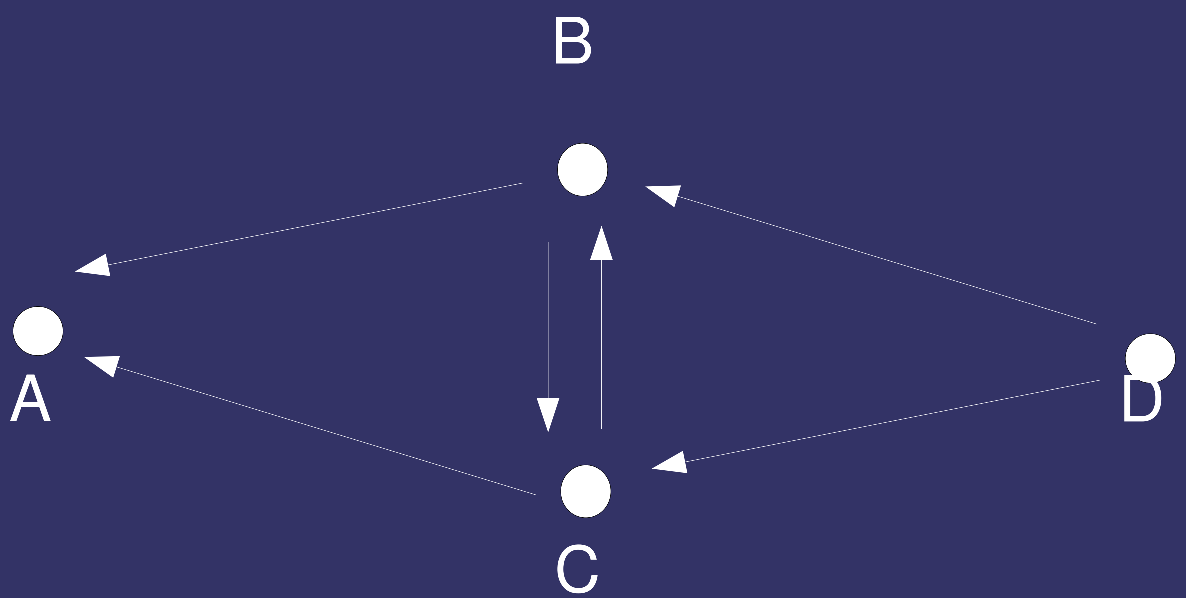
Applying the Evaluation Postulate (1/3)



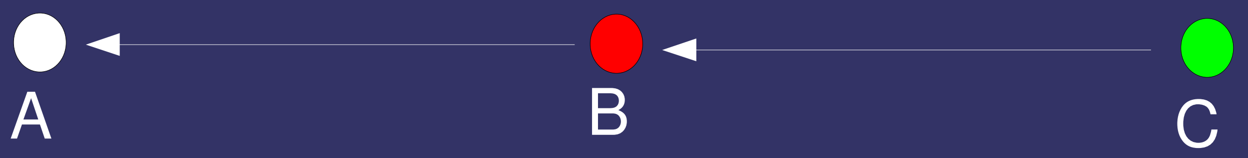


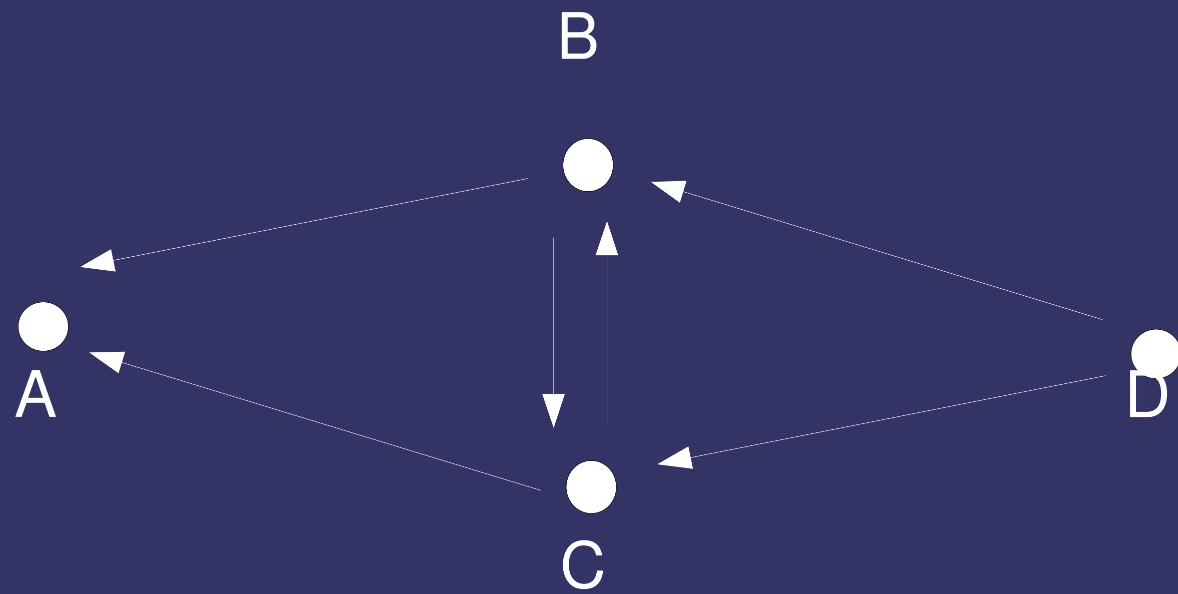
Applying the Evaluation Postulate (1/3)



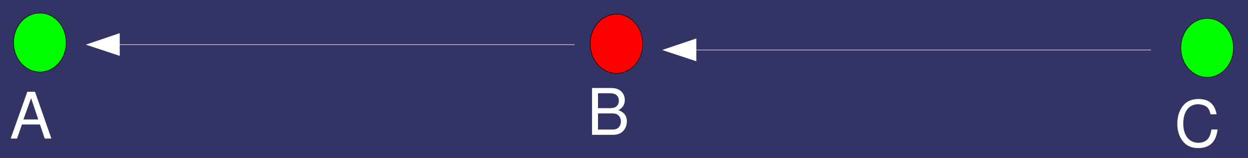


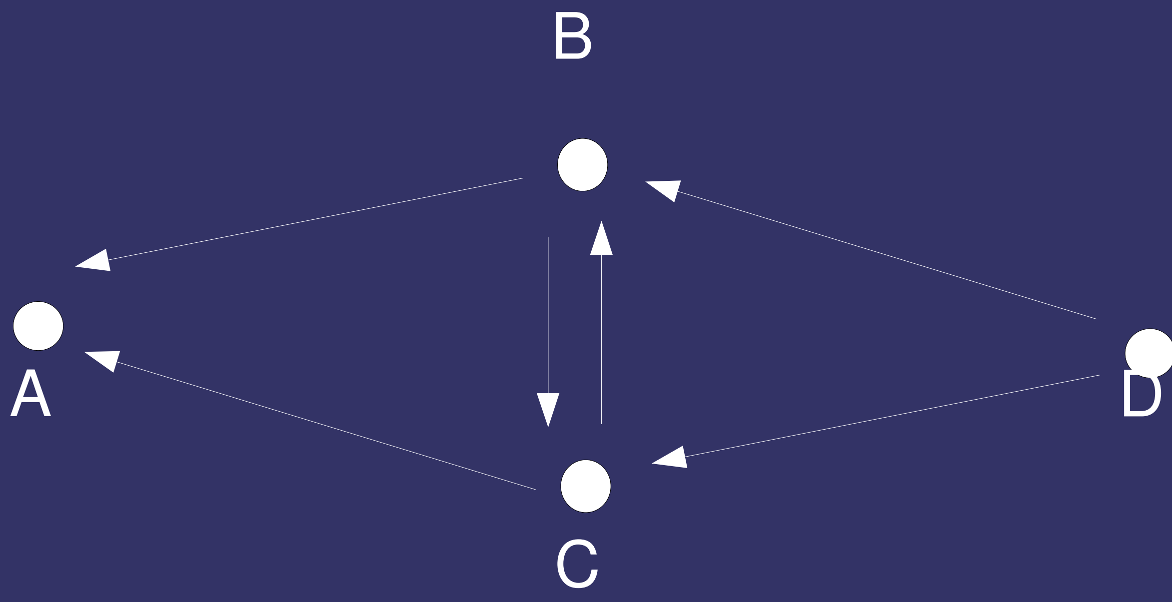
Applying the Evaluation Postulate (1/3)



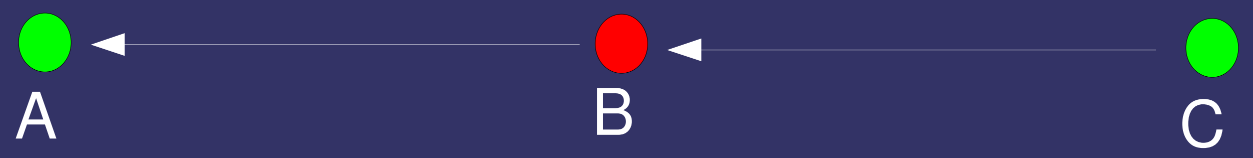


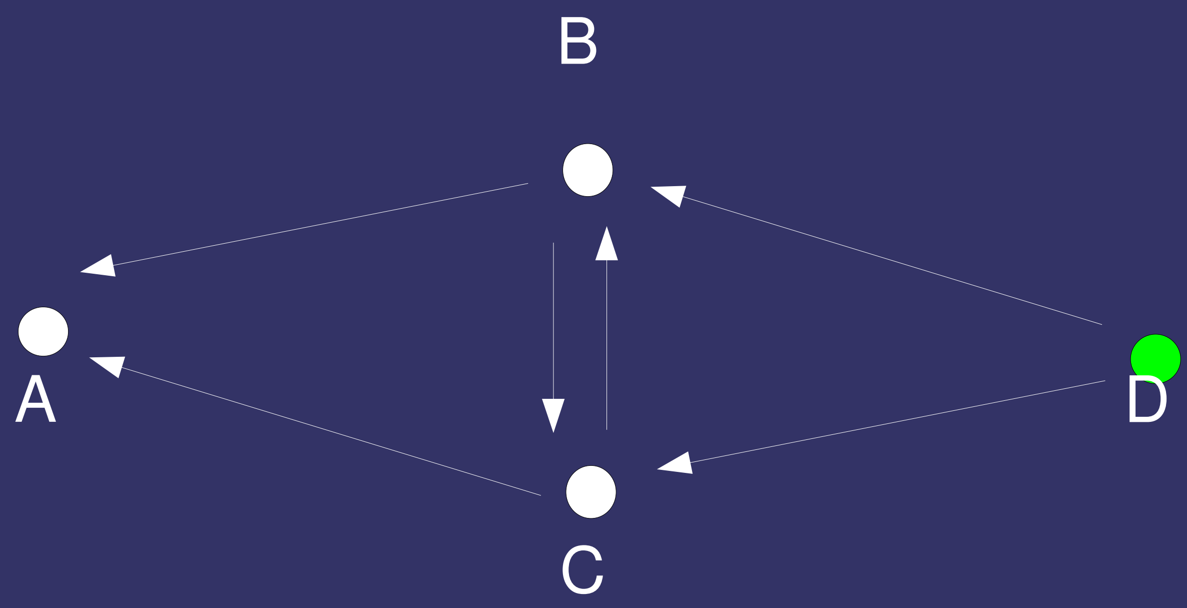
Applying the Evaluation Postulate (1/3)



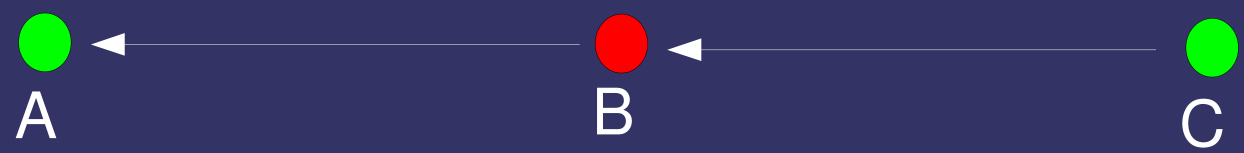


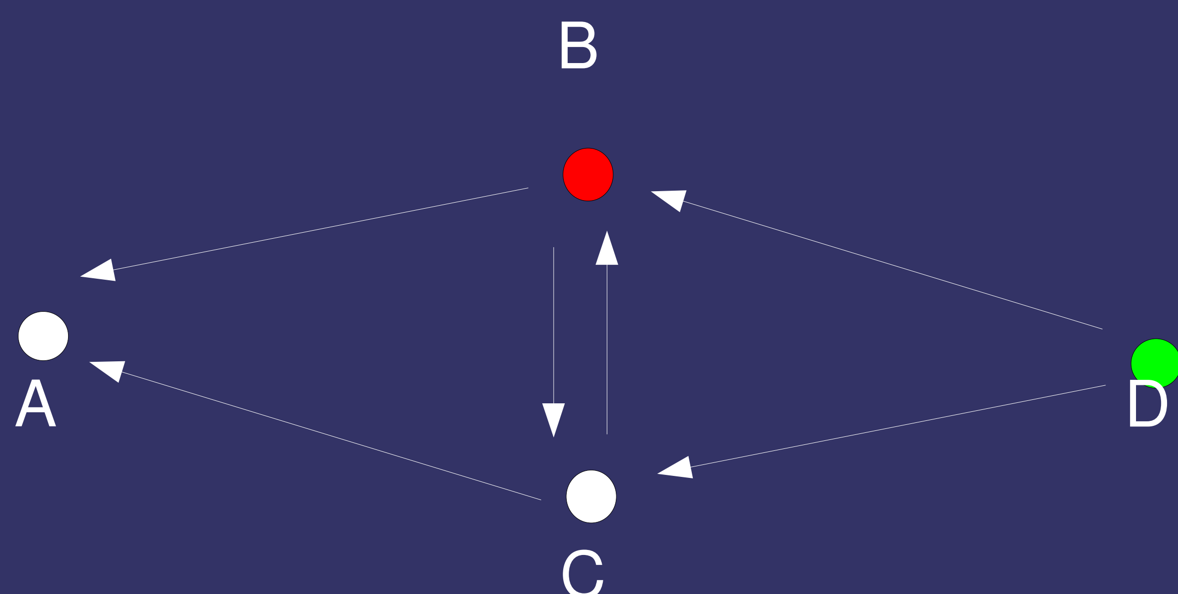
Applying the Evaluation Postulate (1/3)



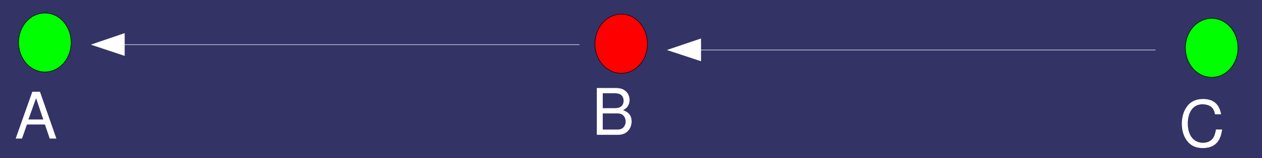


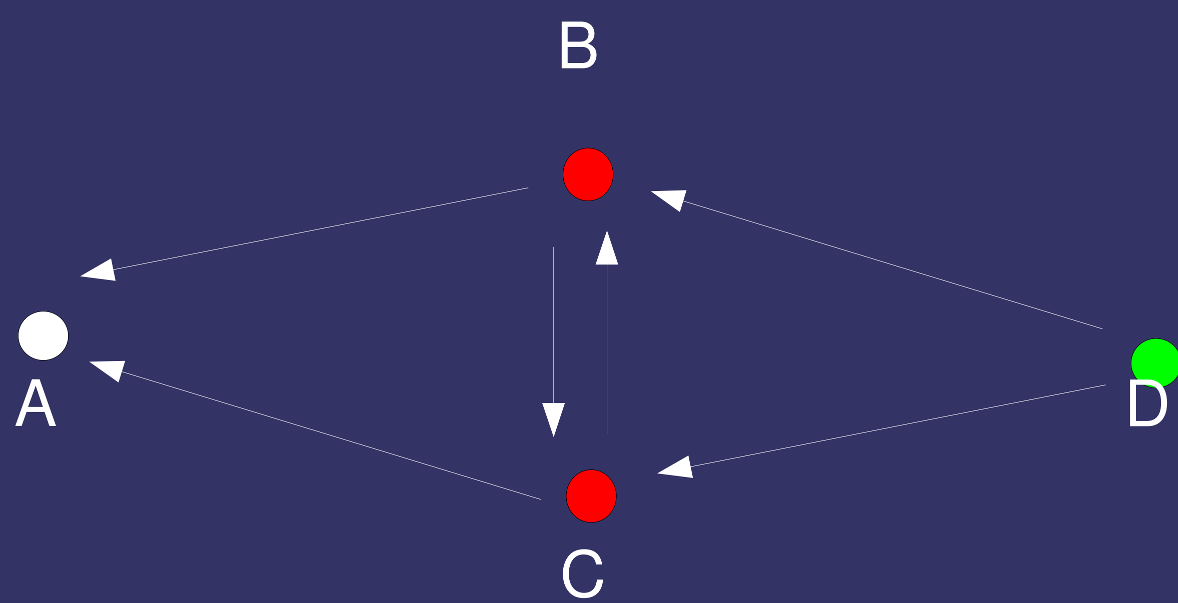
Applying the Evaluation Postulate (1/3)



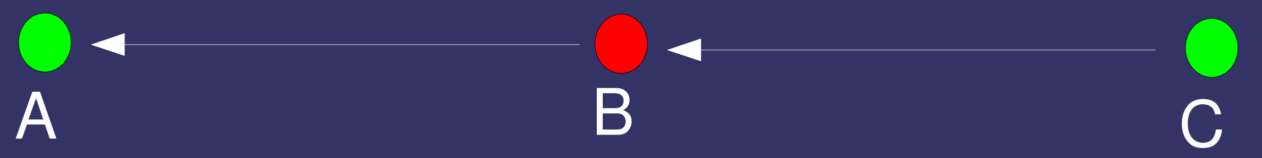


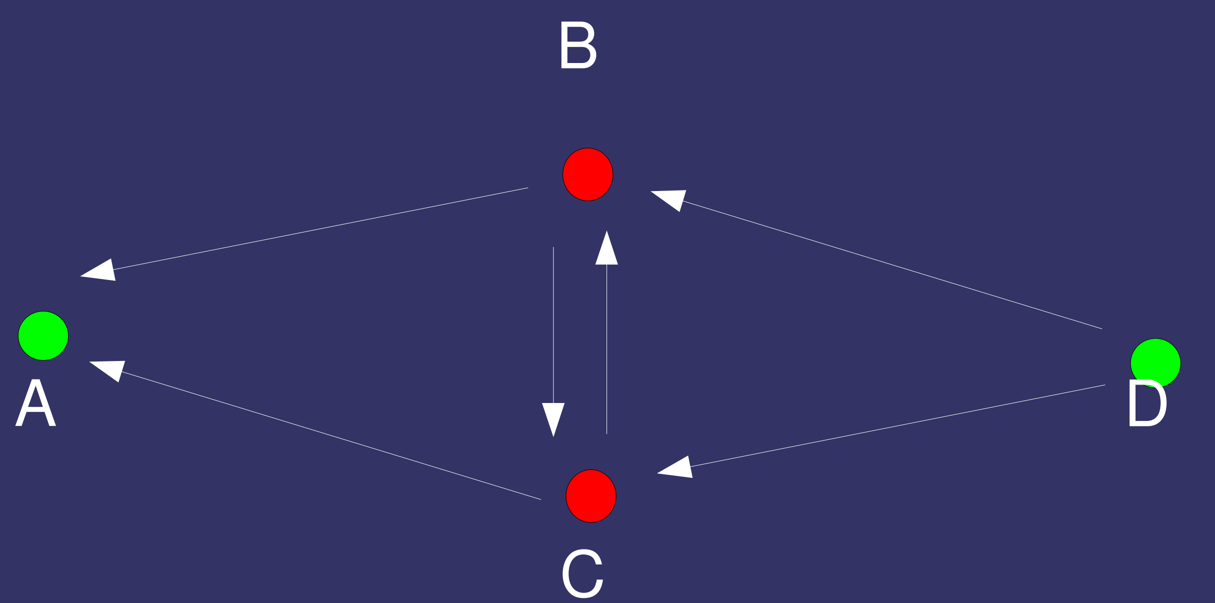
Applying the Evaluation Postulate (1/3)



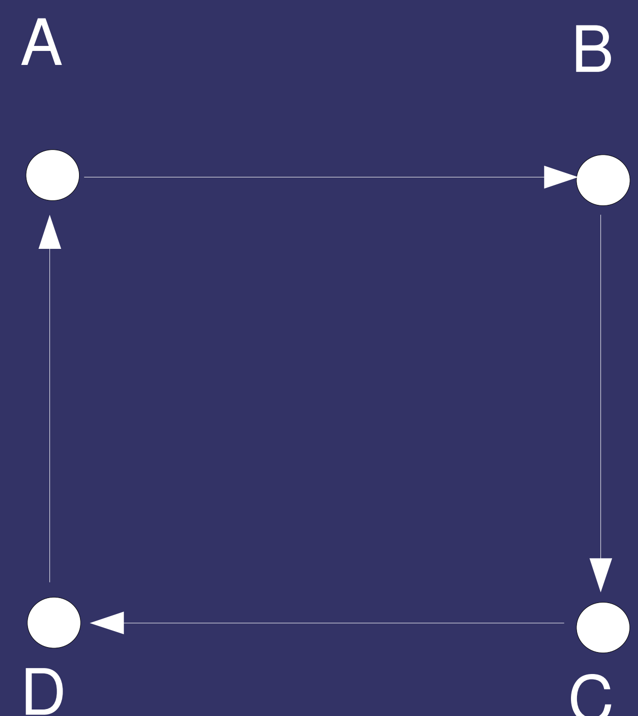


Applying the Evaluation Postulate (1/3)

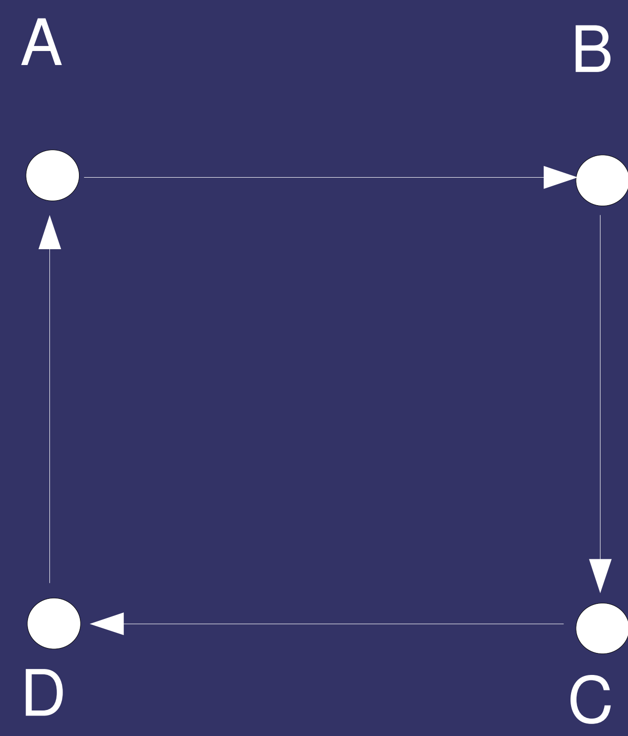




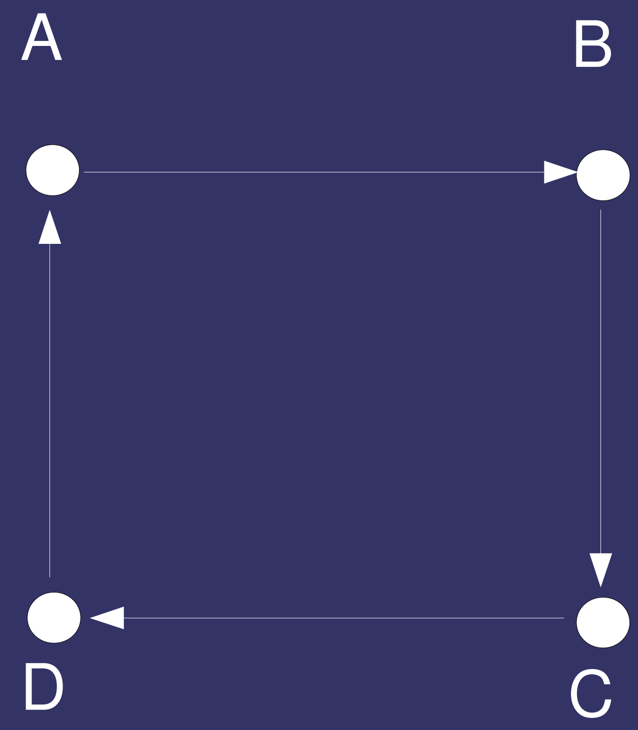
Applying the Evaluation Postulate (2/3)



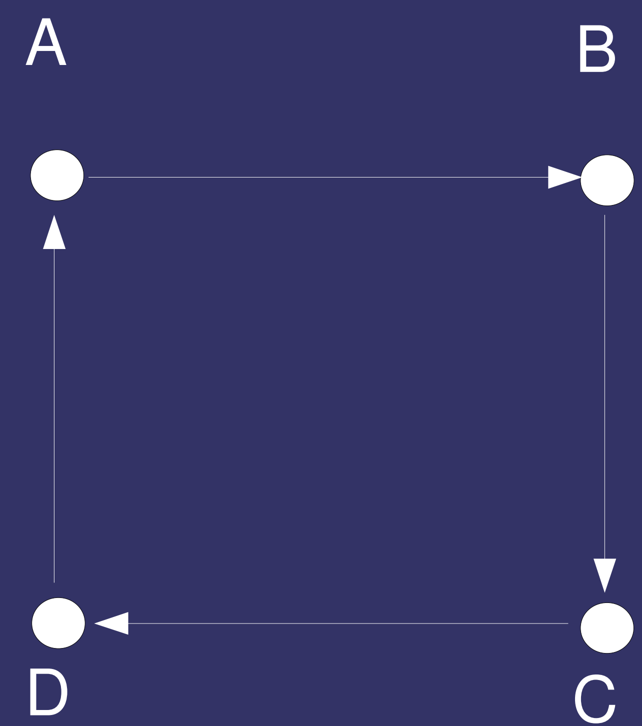
Applying the Evaluation Postulate (2/3)



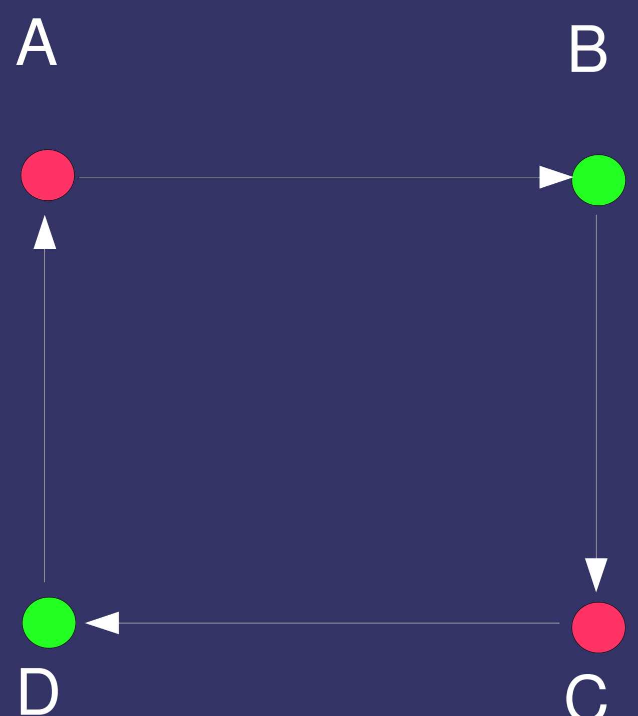
Applying the Evaluation Postulate (2/3)



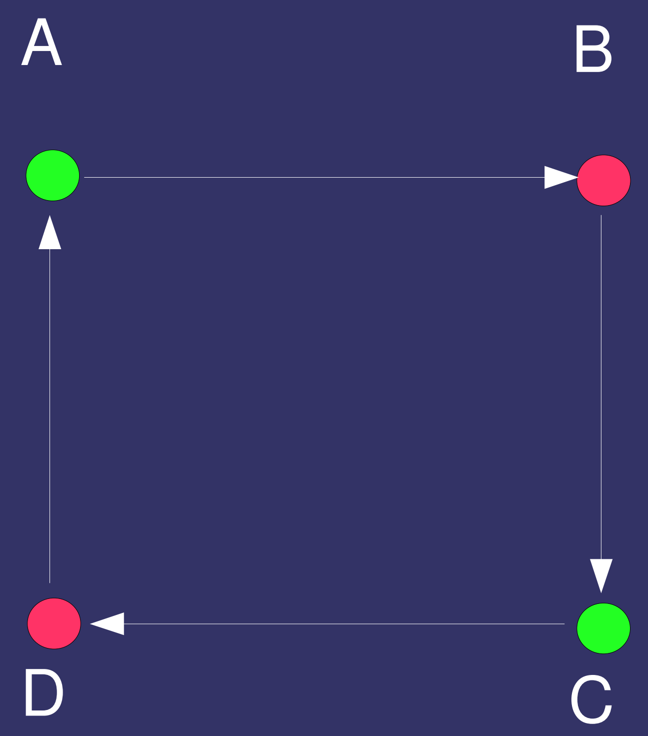
Applying the Evaluation Postulate (2/3)



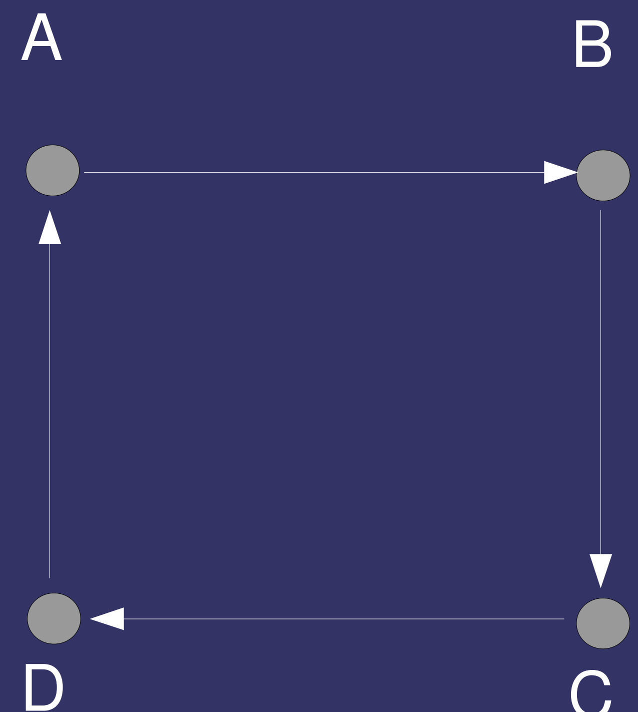
Applying the Evaluation Postulate (2/3)



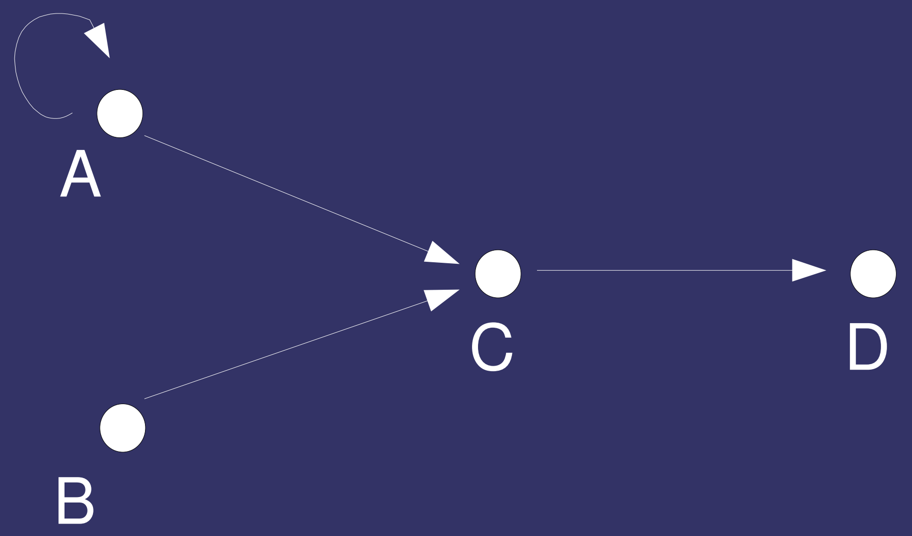
Applying the Evaluation Postulate (2/3)

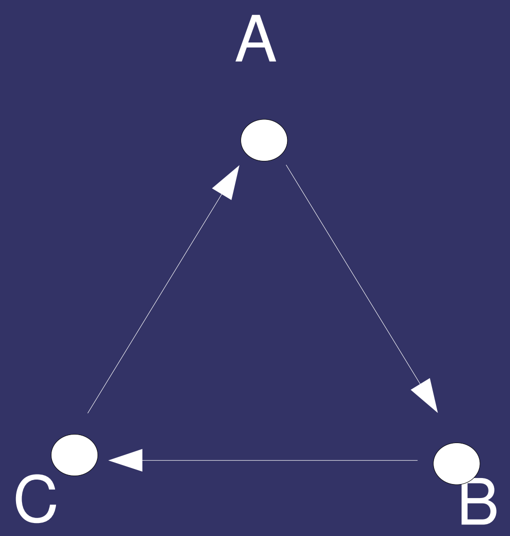


Applying the Evaluation Postulate (2/3)

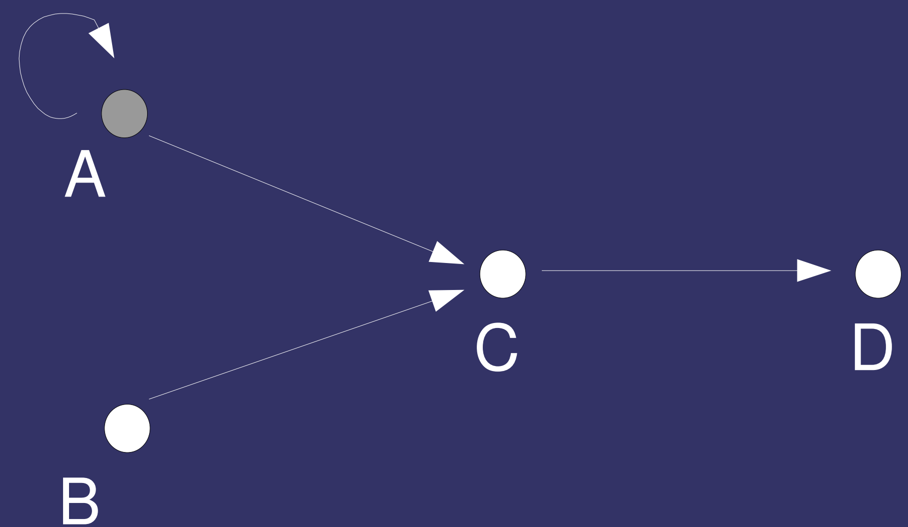


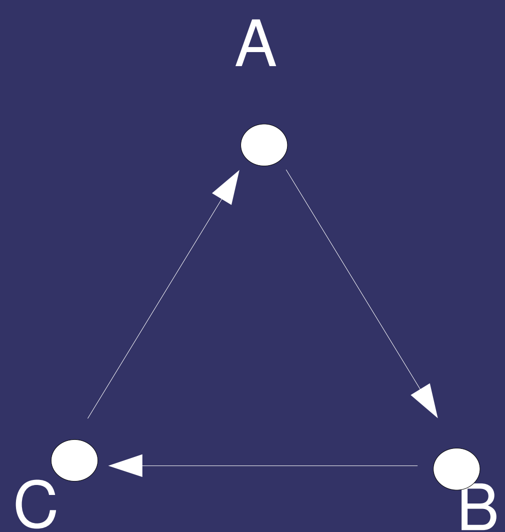
Applying the Evaluation Postulate (3/3)



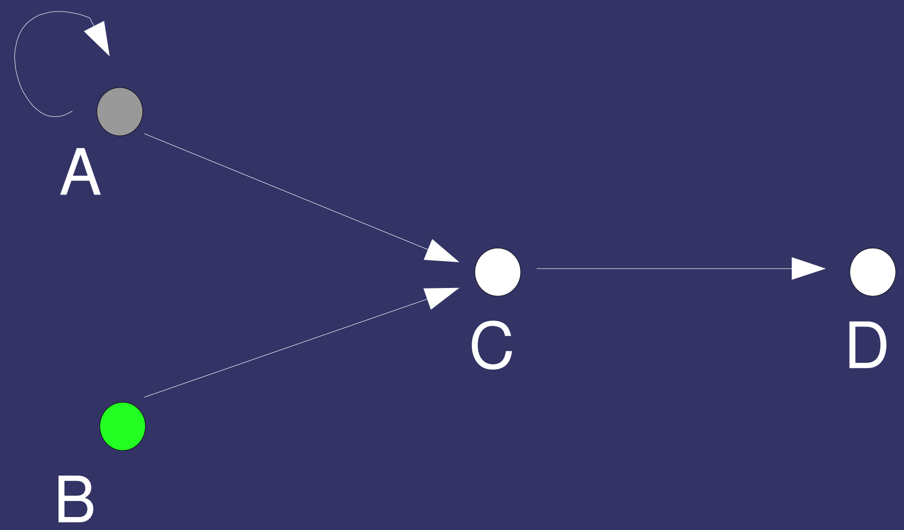


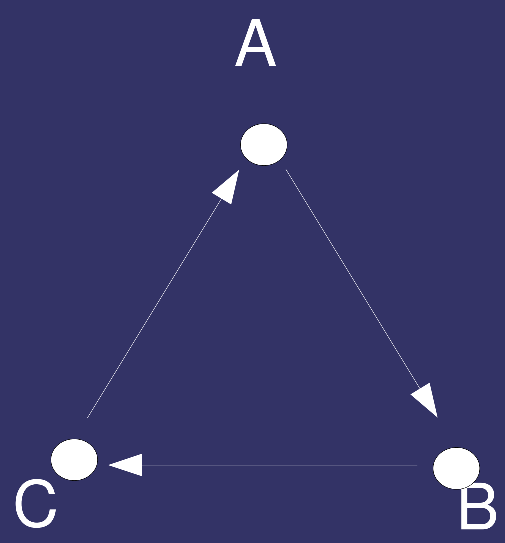
Applying the Evaluation Postulate (3/3)



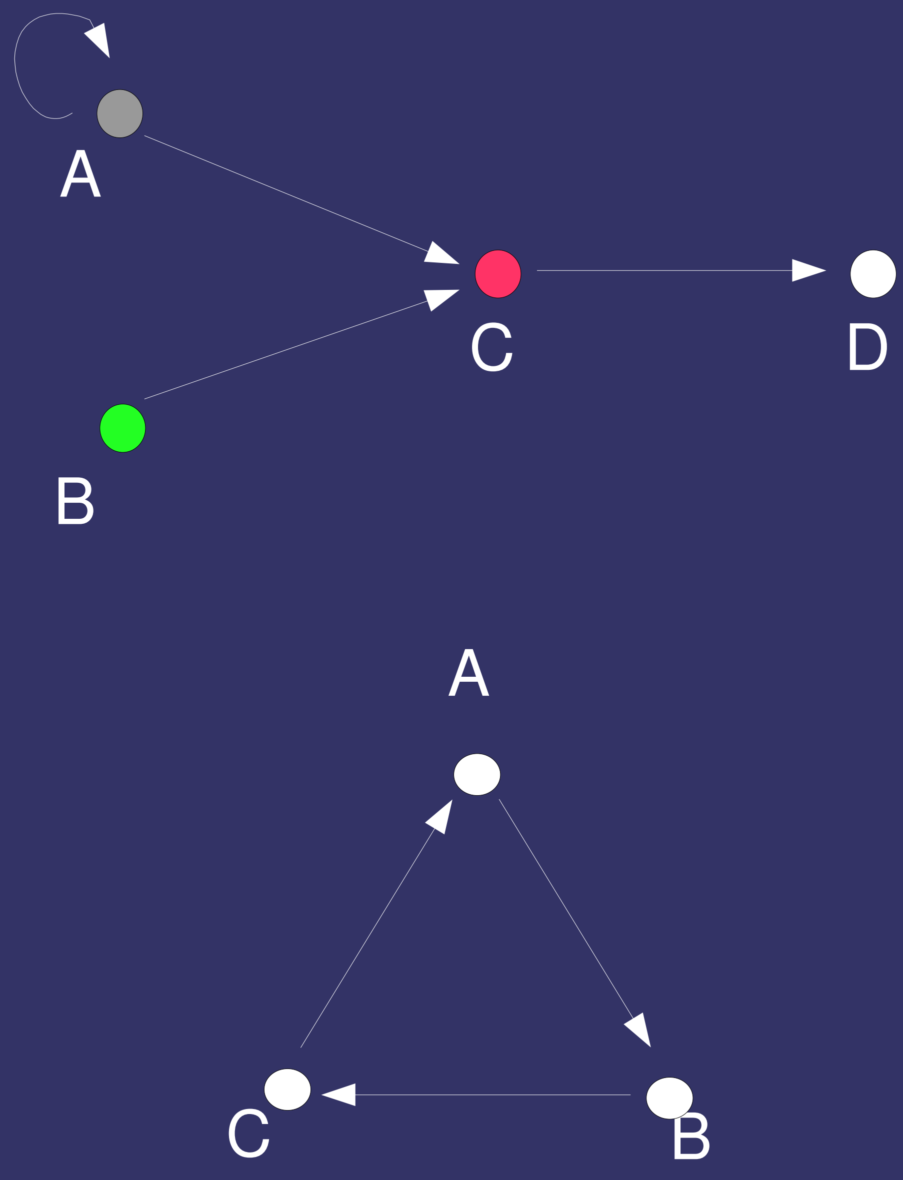


Applying the Evaluation Postulate (3/3)

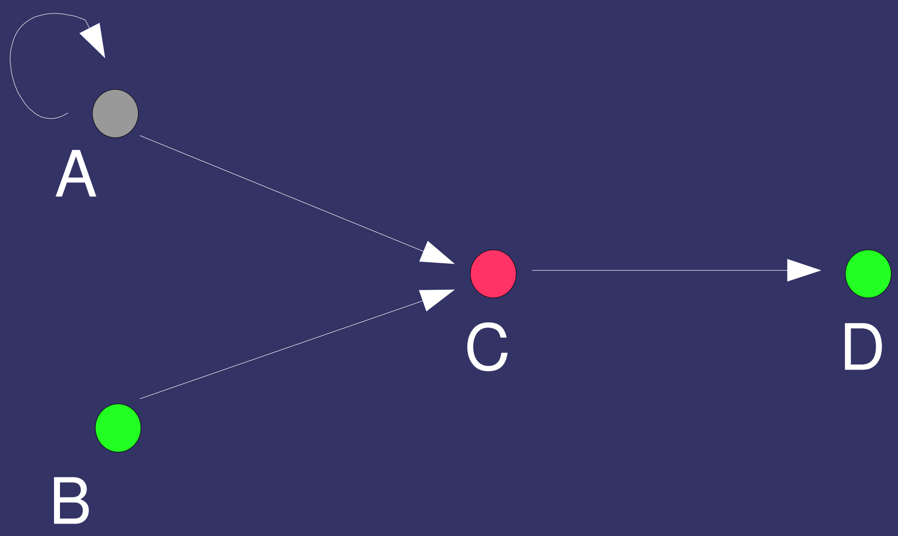


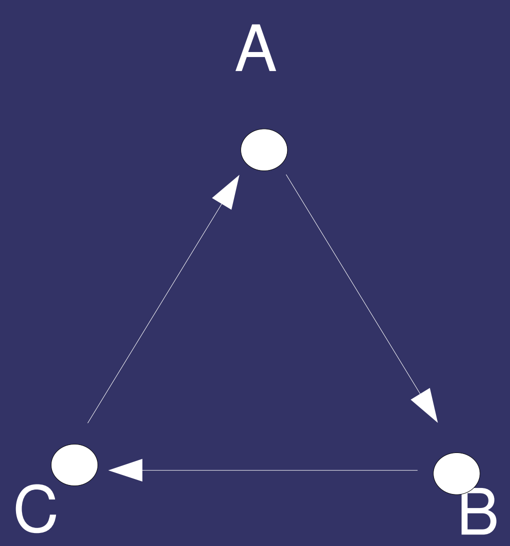


Applying the Evaluation Postulate (3/3)

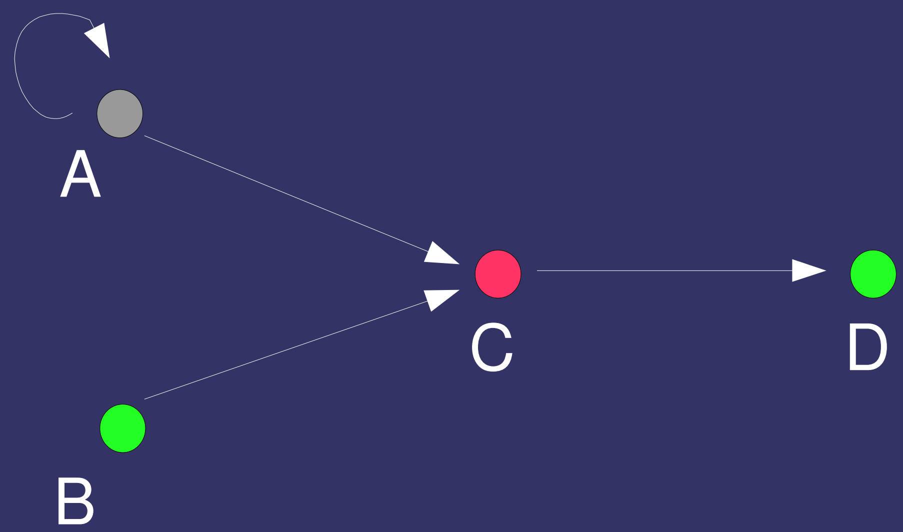


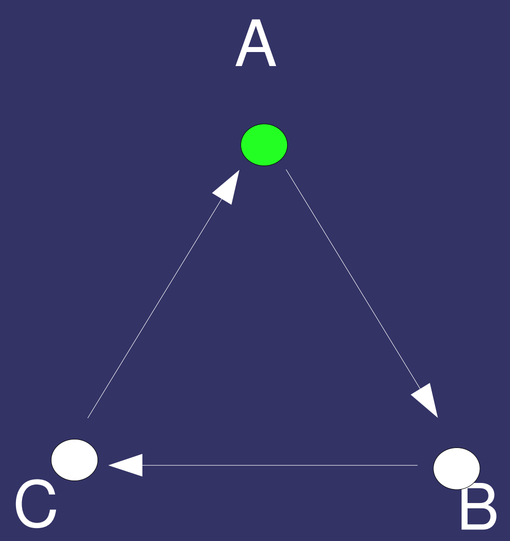
Applying the Evaluation Postulate (3/3)



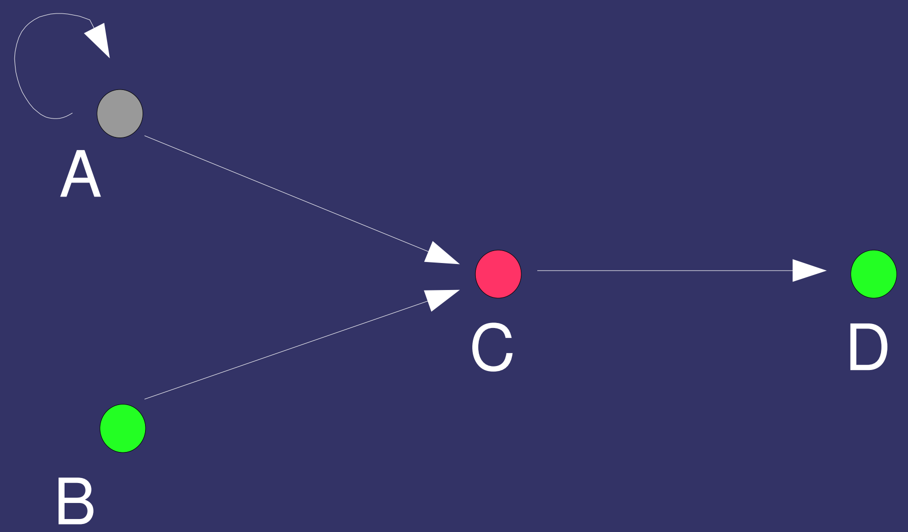


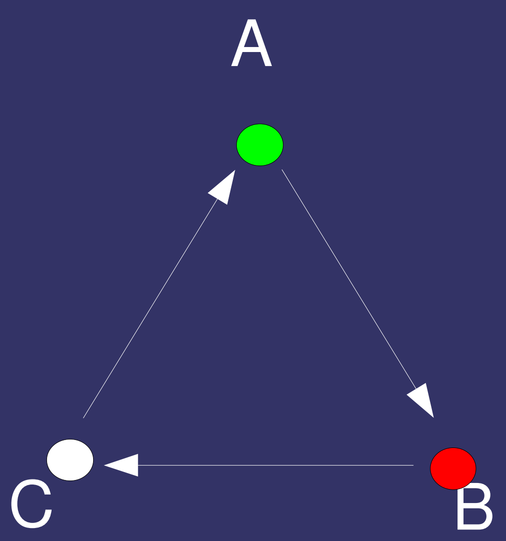
Applying the Evaluation Postulate (3/3)



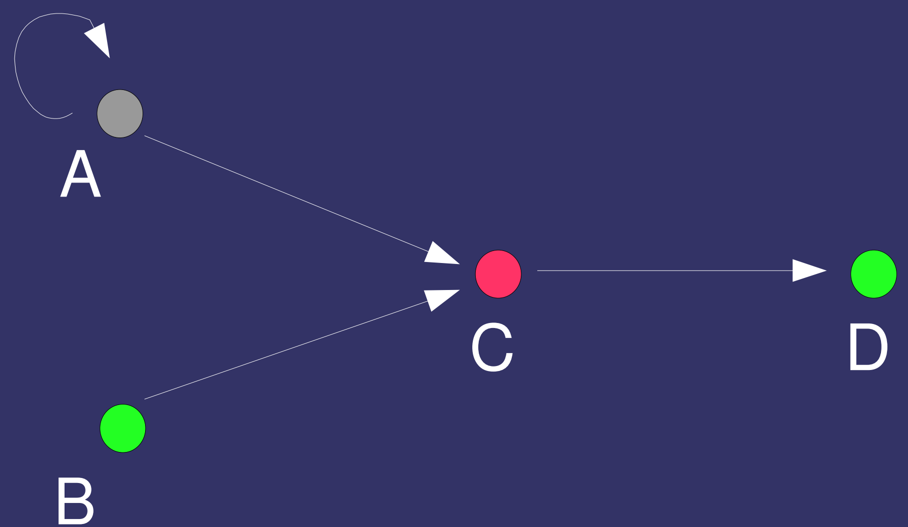


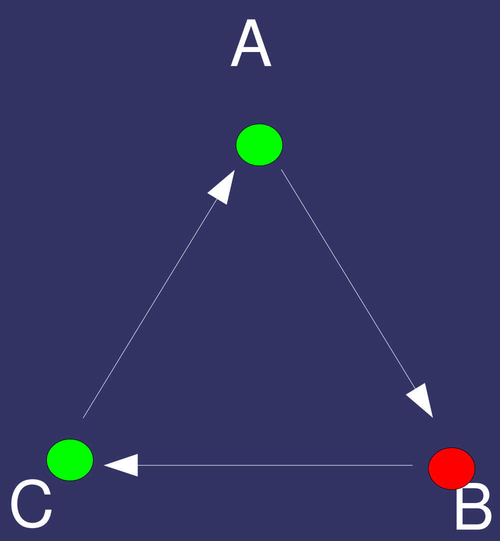
Applying the Evaluation Postulate (3/3)



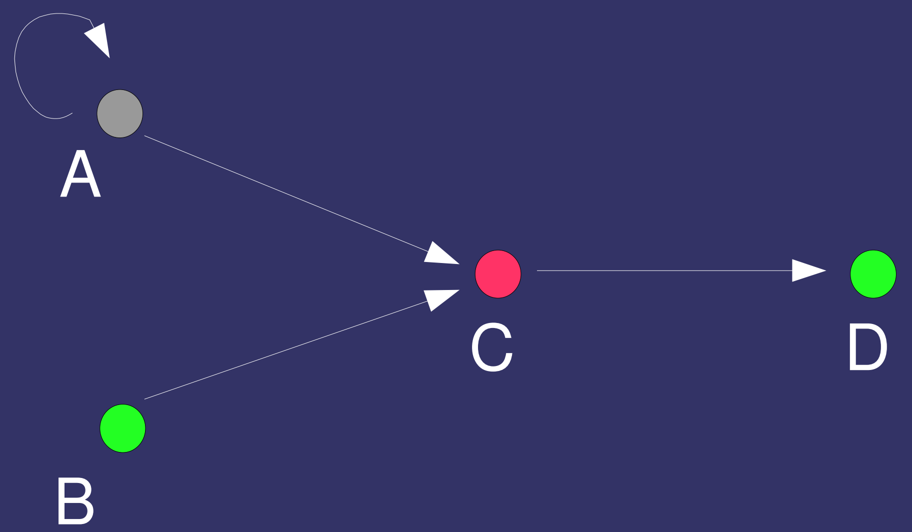


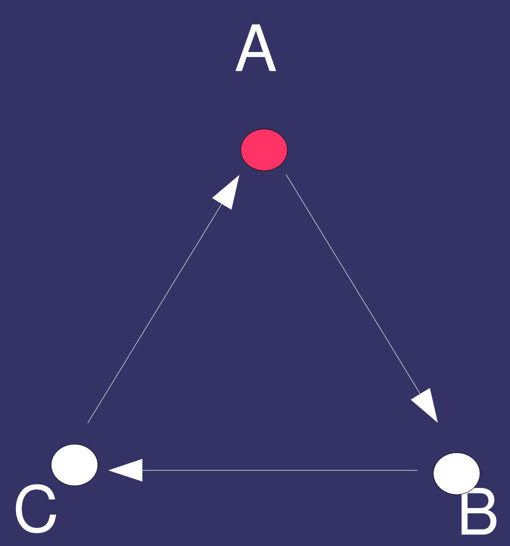
Applying the Evaluation Postulate (3/3)



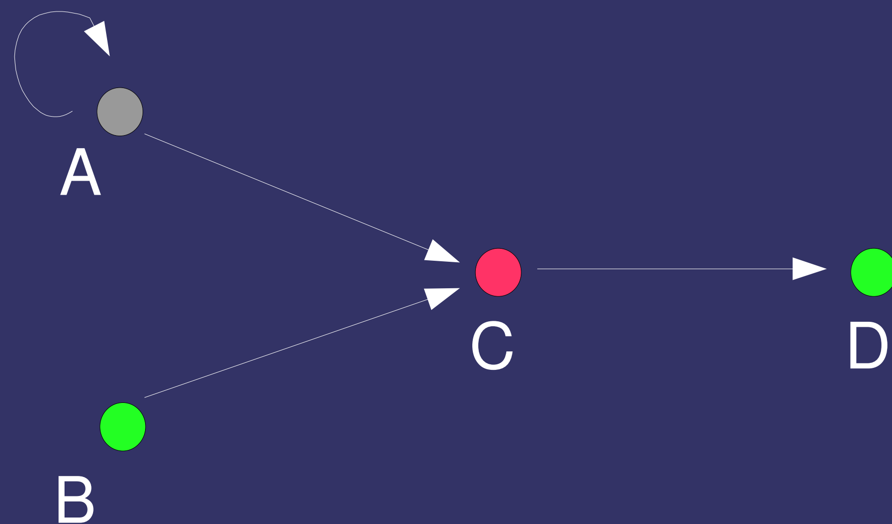


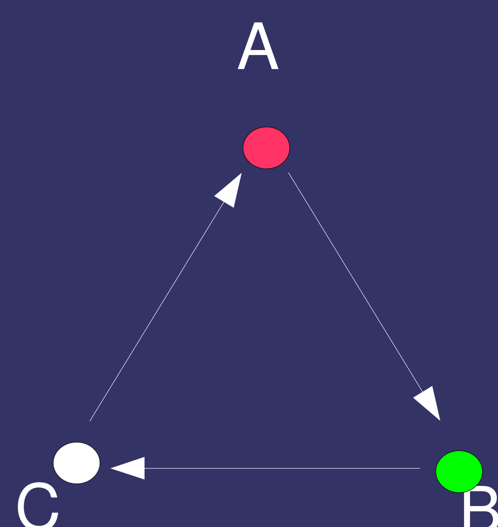
Applying the Evaluation Postulate (3/3)



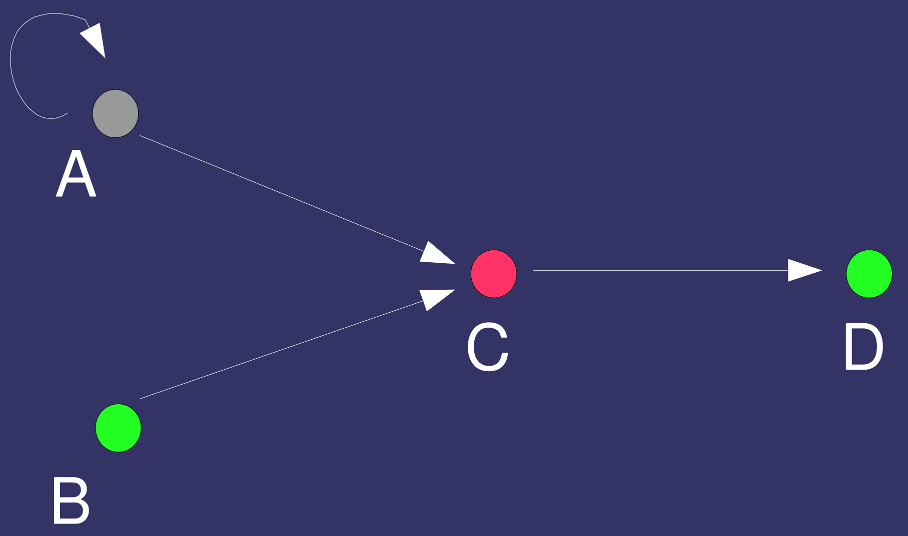


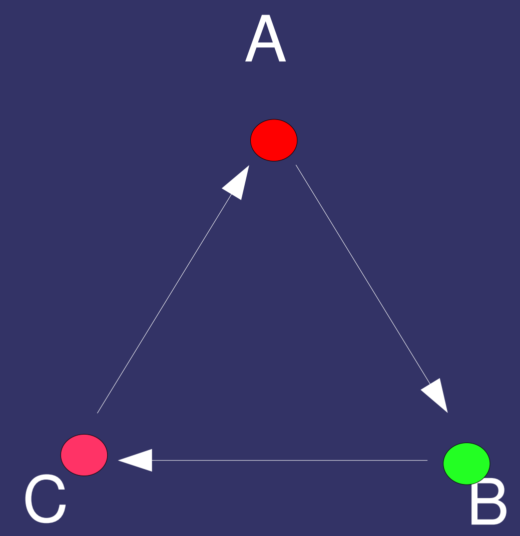
Applying the Evaluation Postulate (3/3)



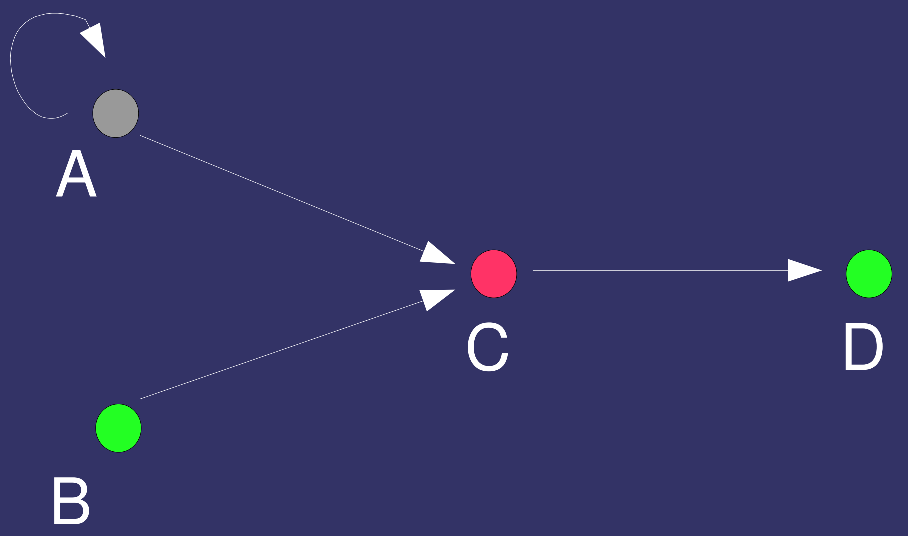


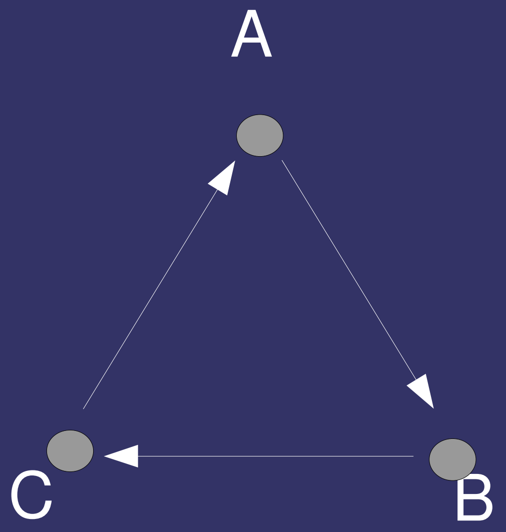
Applying the Evaluation Postulate (3/3)



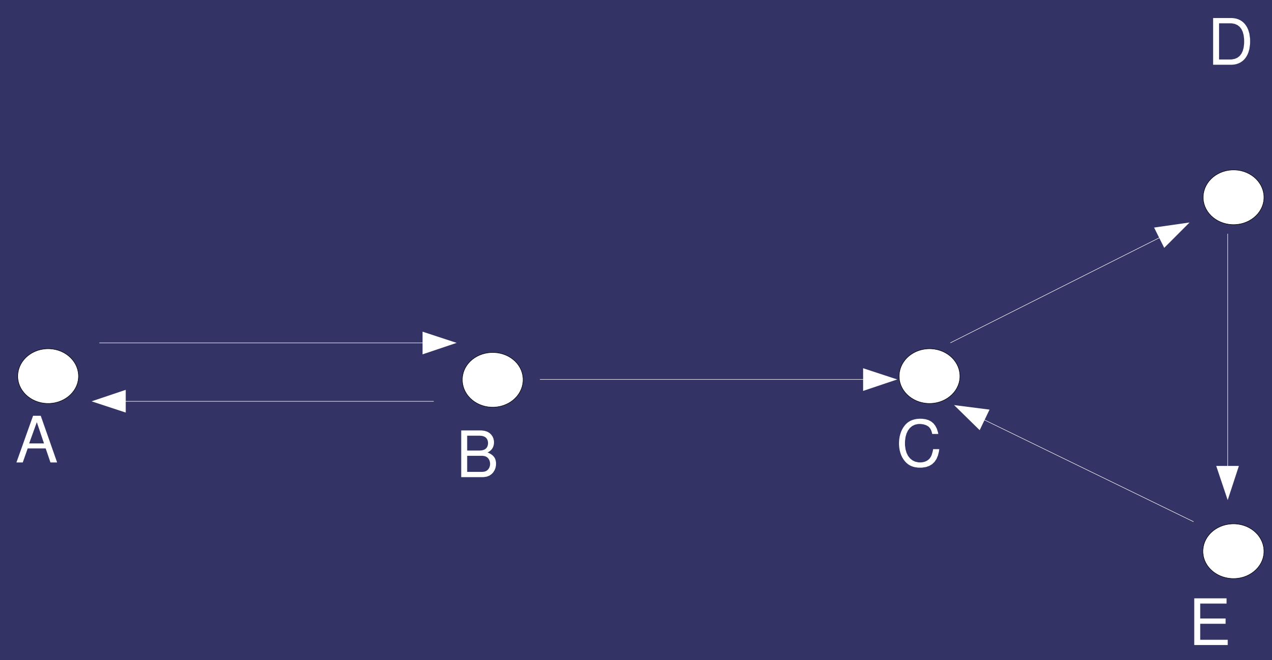


Applying the Evaluation Postulate (3/3)



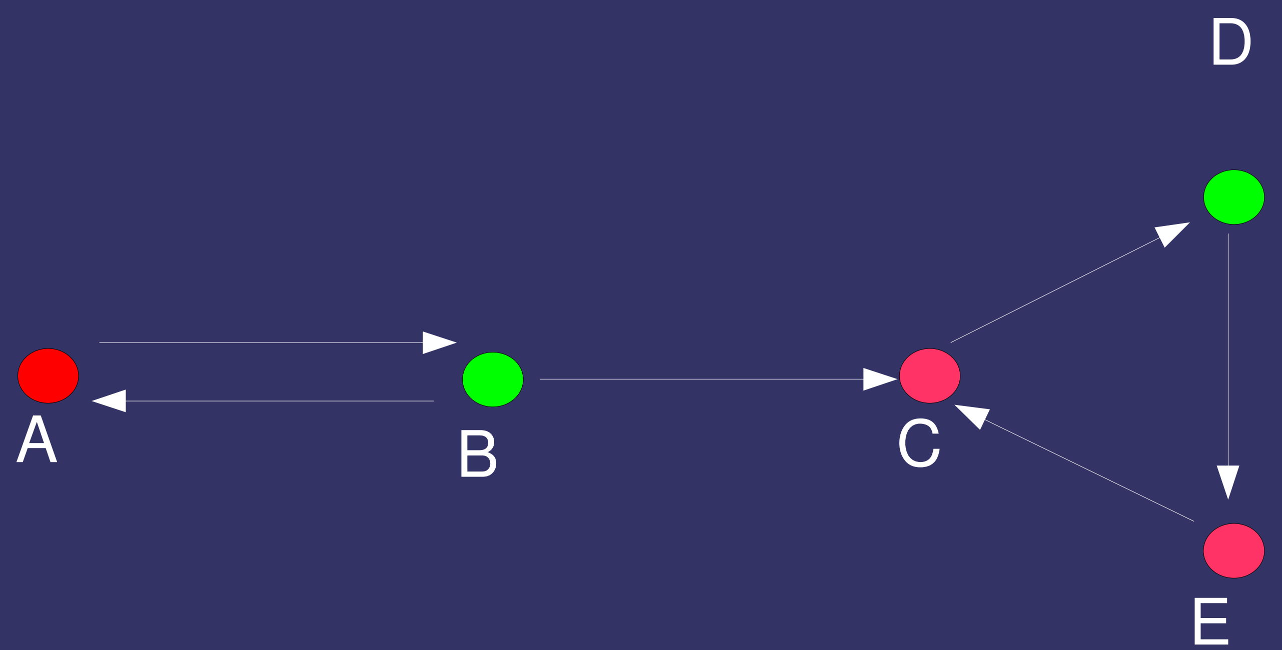


Exercise 1



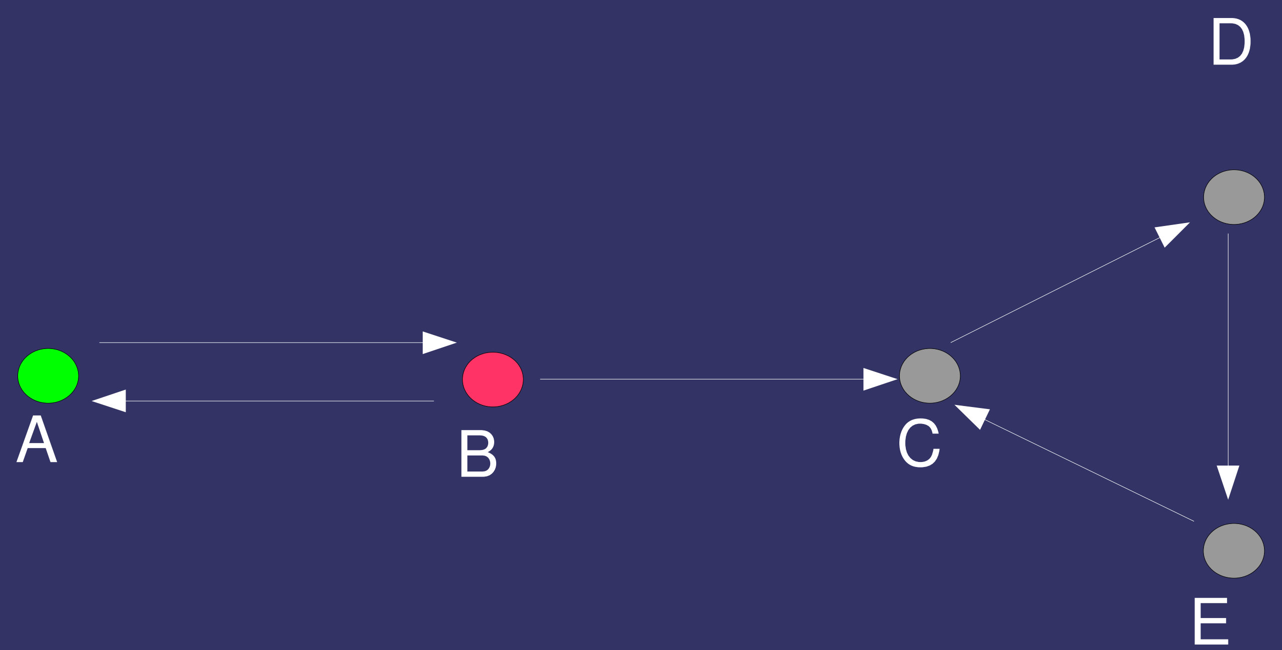
Give the three labellings of this argumentation framework

Exercise 1



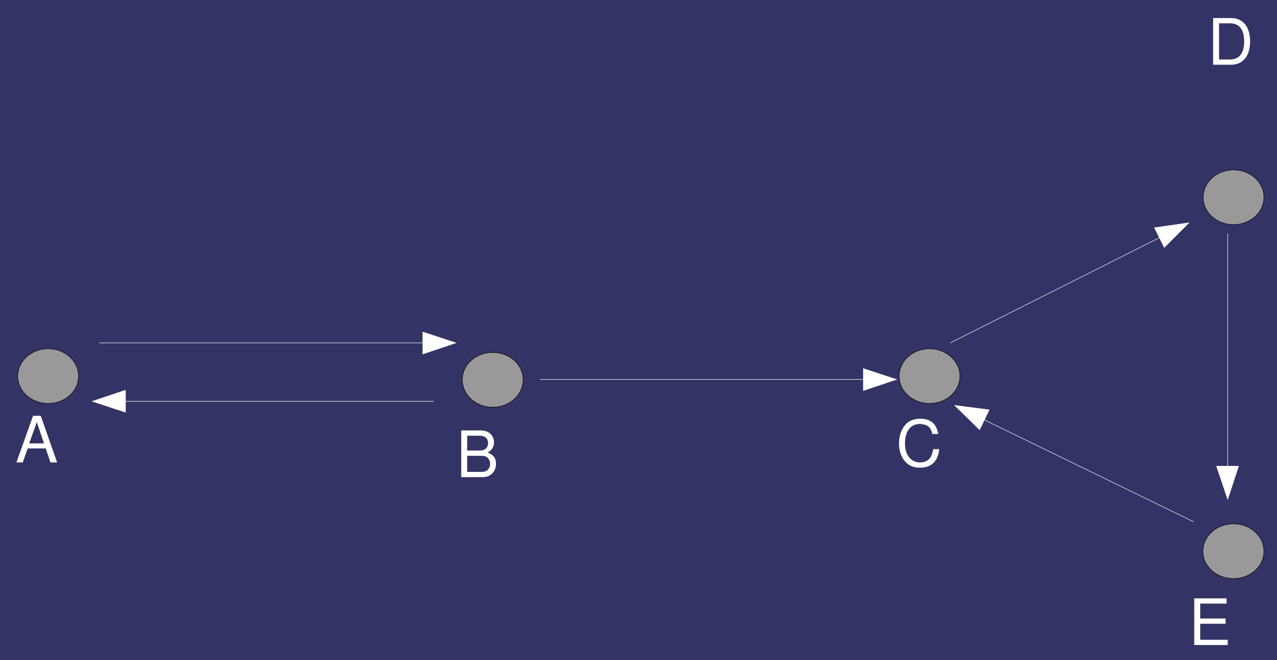
Give the three labellings of this argumentation framework

Exercise 1



Give the three labellings of this argumentation framework

Exercise 1



Give the three labellings of this argumentation framework

Argument Evaluation in the Literature (1/3)

Args is <u>conflict-free</u> iff
Args does not contain A,B such that A defeats B

Args <u>defends</u> an argument A iff
for each argument B that defeats A,
Args contains an argument (C) that defeats B

F(Args) = all arguments defended by Args

Argument Evaluation in the Literature (2/3)

A conflict-free set of arguments Args is called:

<u>admissible</u> iff Args $\subseteq$ F(Args)

a <u>complete extension</u> iff Args = F(Args)

a <u>grounded extension</u> iff Args is the minimal complete extension

a <u>preferred extension</u> iff Args is a maximal admissible set

a <u>stable extension</u> iff Args is a conflict-free set that defeats everything not in it

a <u>semi-stable extension</u> iff Args is an admissible set with Args $\cup$ Args$^+$ maximal

Argument Evaluation in the Literature (3/3)

A conflict-free set of arguments Args is called:

admissible iff Args $\subseteq$ F(Args)

a complete extension iff
Args = F(Args)

a grounded extension iff
Args is the minimal complete extension

a preferred extension iff
Args is a maximal complete extension

a stable extension iff Args is a complete extension that defeats everything not in it

a semi-stable extension iff Args is a complete extension with Args $\cup$ Args$^+$ maximal

Literature and Labellings

restriction on compl. labeling	***Dung-style semantics***
no restrictions	complete semantics
empty undec	stable semantics
maximal in	preferred semantics
maximal out	preferred semantics
maximal undec	grounded semantics
minimal in	grounded semantics
minimal out	grounded semantics
minimal undec	semi-stable semantics

Some properties of argument semantics

grounded extension = $\cap$ complete extensions
[Dung 1995 AIJ]

an argument is in at least one preferred extension
iff it is in at least one complete extension
iff it is in at least one admissible set.

Computing the Grounded Extension

Idea: start with the undefeated arguments,
then iteratively add the defended arguments

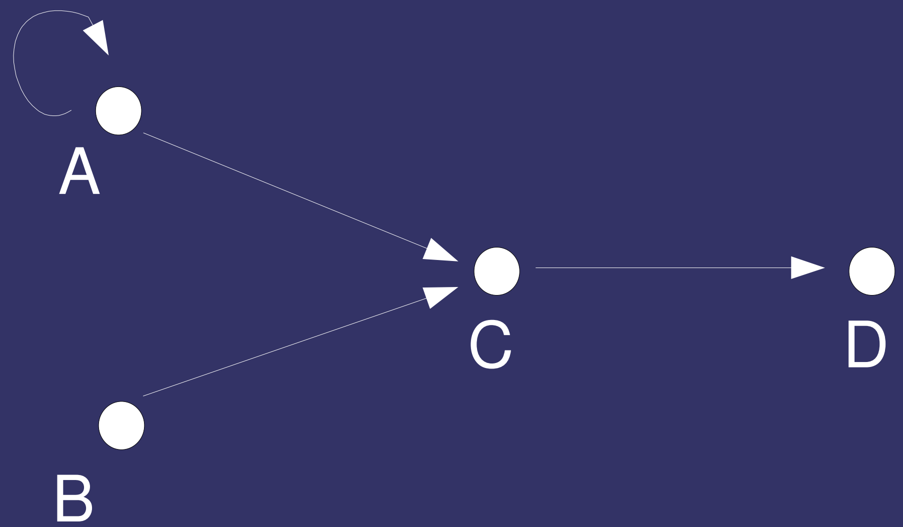
$F^0 = \emptyset$

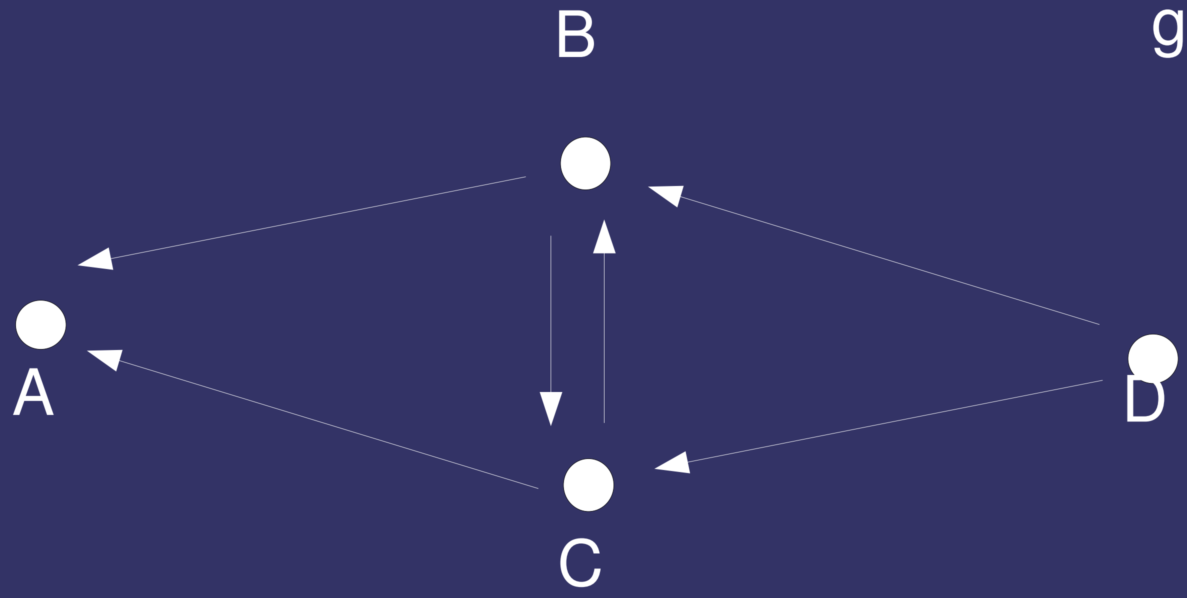
$F^{i+1} = \{ A \mid A \text{ is defended by } F^i \}$

$F^\infty = \bigcup_{i=0\ldots\infty} F^i$

If each argument has a finite set of defeaters,
then F^∞ is the grounded extension.

Exercise 2





Give for each of these argumentation frameworks the grounded extension

A Dialectical Game for Grounded Semantics

Is argument A element of the grounded extension?

proponent states A

opponent and proponent then take turns, in which they state an argument thats defeat the previous argument

proponent is not allowed to repeat any previous argument

a player wins iff the other player cannot move

Argument A is in the grounded extension iff proponent has winning strategy for A

A Dialectical Game for Admissibility

Is argument A element of an admissible set?

proponent states A

opponent and proponent then take turns;

the opponent each time states an argument that defeats *one of the previous arguments* of the proponent; the proponent each time states an argument that defeats *the immediately preceding argument* of the opponent

the proponent may repeat its own moves,

but not the moves of the opponent;

the opponent may repeat the proponent's moves

but not its own moves

proponent wins iff opponent cannot move;

opponent wins iff proponent cannot move

or if opponent is able to repeat proponent's move

A is in admissible set iff proponent can win game

Default Logic as Argumentation

default: pre(d): jus(d) / cons(d)

arguments of the form: $(d_1, ..., d_n)$ where

for each d_i $(1 \leq i \leq n)$ it holds that

$\{cons(d_1), ..., cons(d_{i-1})\} \cup W \vdash pre(d_i)$

$(d_1, ..., d_n)$ defeats $(d'_1, ..., d'_m)$ iff

there is some d'_i $(1 \leq i \leq m)$ such that

$\{cons(d_1), ..., cons(d_n)\} \cup W \vdash \neg jus(d'_i)$

stable semantics

Pollock

Arguments of the form ($pfrule_1$, ..., $pfrule_n$) where $\{cons(pfrule_1), ..., cons(pfrule_{i-1})\} \cup W \vdash ant(pfrule_i)$

defeat: rebutting + undercutting

preferred semantics (before: grounded semantics)

Logic Programming

Arguments: trees constructed with rules. The children of a rule
$c \leftarrow a_1, ..., a_n$, not $b_1, ..., $ not b_m
are rules with heads $a_1, ..., a_n$
An argument A defeats an argument B iff
A contains a rule with c as its head and
B contains a rule with not c in its body
stable semantics (“stable model semantics”)
grounded semantics (“well-founded semantics”)

How Things Go Wrong (1/5)

r	$r \Rightarrow m$	$m \rightarrow hs$
p	$p \Rightarrow b$	$b \rightarrow \neg hs$

$A1 = (r) \Rightarrow m$

$A2 = (p) \Rightarrow b$

$A3 = A1 \rightarrow hs$

$A4 = A2 \rightarrow \neg hs$

Conclusions m and b are justified under any semantics but what about hs and $\neg$hs?

How Things Go Wrong (2/5)

r	$r \Rightarrow m$	$m \supset hs$	(“$\rightarrow$”$\equiv$“$\vdash$”)
p	$p \Rightarrow b$	$b \supset \neg hs$	

A1: $(r) \Rightarrow m$

A2: $(p) \Rightarrow b$

A3: $(A1,\ m \supset hs) \rightarrow hs$

A4: $(A2,\ b \supset \neg hs) \rightarrow \neg hs$

A5: $(A3,\ b \supset \neg hs) \rightarrow \neg b$

A6: $(A4,\ m \supset hs) \rightarrow \neg m$

So far,
so good...

How Things Go Wrong (3/5)

$j \qquad j \Rightarrow s$

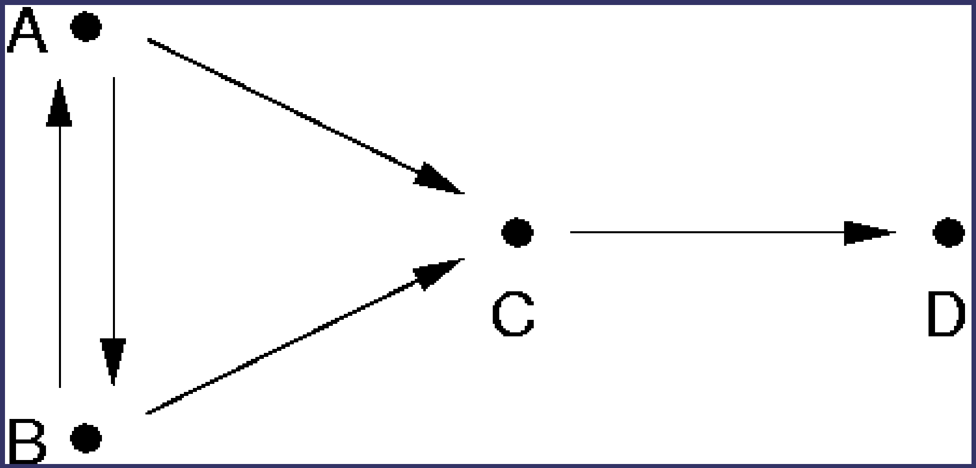
$(\text{“}\rightarrow\text{”} \equiv \text{“}\vdash\text{”})\ m \qquad m \Rightarrow \neg s \quad wf \quad wf \Rightarrow r$

There now exist the following arguments:

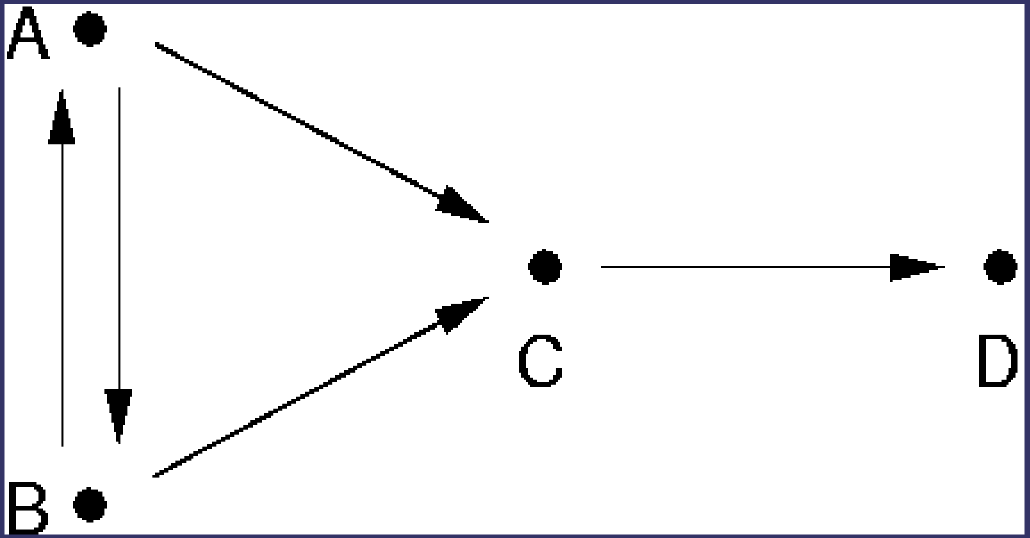
$A = (j) \Rightarrow s$ (unfortunately,

$B = (m) \Rightarrow \neg s$ there also exists:

$D = (wf) \Rightarrow r$ $C = A, B \rightarrow \neg r)$



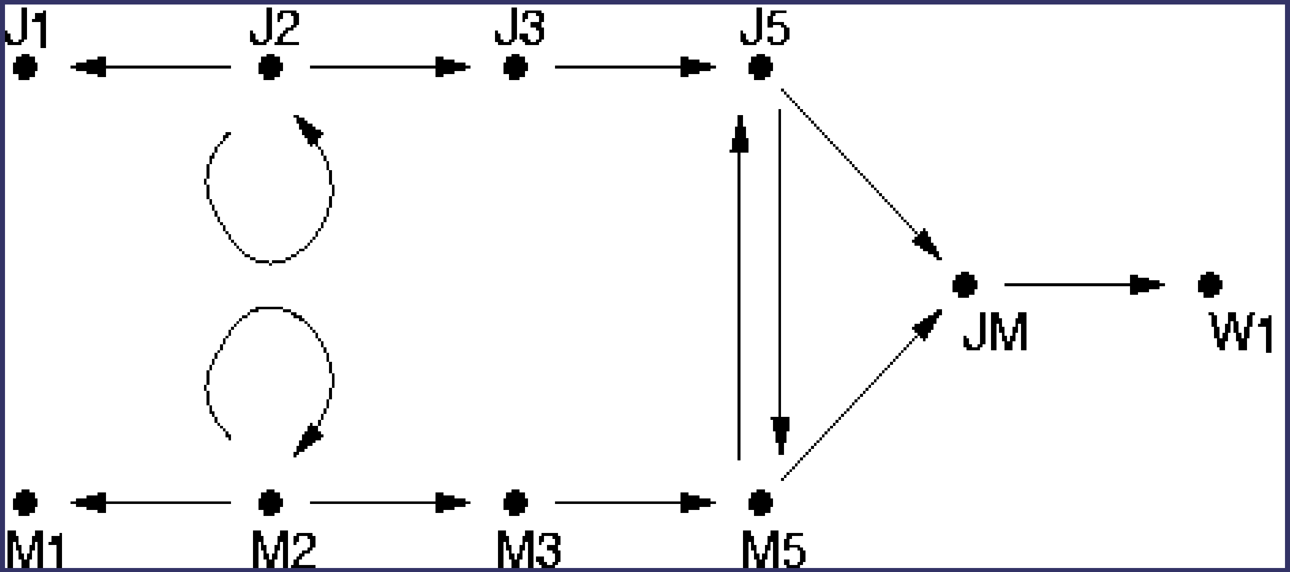
How Things Go Wrong (4/5)



Grounded semantics: no justified arguments
Why not use preferred or stable semantics?
Reiter and Pollock also do this...

How Things Go Wrong (5/5)

John: “Cup of coffee contains sugar.”
Mary: “Cup of coffee doesn't contain sugar.”
John: “I'm unreliable.”
Mary: “I'm unreliable.”
Weather Forecaster: “Tomorrow rain.”



Quality Postulates

Let J be the justified conclusions and
$Cl_S(J)$ be the closure of J under the rules in S.
direct consistency: $\neg\exists p: (p \in J \wedge \neg p \in J)$
closedness: $J = Cl_S(J)$
indirect consistency: $Cl_S(J)$ is consistent
crash-resistancy:
"Local problems should not have global effects"
non-interference:
"A set of formulas should not be able to influence the entailment of a totally unrelated set of formulas, when being merged to it"

Transposition

Let *s* be a strict rule of the form

$a_1, ..., a_n \rightarrow c$

A rule *s´* is a *transposition* of *s* iff *s´* is of the form

$a_1, ..., a_{i-1}, \neg c, a_{i+1}, ..., a_n \rightarrow \neg a_i$

(for some $1 \leq i \leq n$)

A set of strict rules S is closed under transposition iff for each $s \in S$, if s' is a transposition of *s* then $s' \in S$.

Restricted versus unrestricted rebut

$((a) \Rightarrow b) \Rightarrow c$

$((d) \Rightarrow e) \Rightarrow \neg c$

Restricted versus unrestricted rebut

$((a) \rightarrow b) \rightarrow c$

$((d) \Rightarrow e) \Rightarrow \neg c$

Restricted versus unrestricted rebut

$((a) \Rightarrow b) \rightarrow c$

$((d) \rightarrow e) \Rightarrow \neg c$

Restricted versus unrestricted rebut

$((a) \Rightarrow b) \rightarrow c$

$((d) \rightarrow e) \Rightarrow \neg c$

unrestricted rebut:
an argument can be rebutted on a conclusion derived by at least one defeasible rule

restricted rebut:
an argument can be rebutted only on the direct consequent of a defeasible rule

Satisfying the quality postulates

Two possibilities:

strict rules closed under transposition
+ unrestricted rebut
+ grounded semantics

strict rules closed under transposition
+ restricted rebut
+ any “well behaved” semantics

With “well behaved” semantics we mean a semantics that yields a non-empty subset of the complete extensions (e.g. preferred, grounded, complete, ideal, semi-stable, ...)

Floating conclusions (1/2)

"Lars has a Dutch mother, so he's probably Dutch, so he probably likes ice-skating"

"Lars has a Norwegian father, so he's probably Norwegian, so he probably likes ice-skating"

A: $((\text{dutch\_mom}) \Rightarrow \text{dutch}) \Rightarrow \text{likes\_skating}$

B: $((\text{norw\_dad}) \Rightarrow \text{norw}) \Rightarrow \text{likes\_skating}$

C: $((\text{dutch\_mom}) \Rightarrow \text{dutch}) \rightarrow \neg\text{norw}$

D: $((\text{norw\_dad}) \Rightarrow \text{norw}) \rightarrow \neg\text{dutch}$

Here, C defeats A and D, and D defeats C and B.

Grounded extension: $\emptyset$

Wanted: *floating conclusions*

Floating conclusions (2/2)

"Witness X says the suspect killed the victim with an axe on Monday morning"

"Witness Y says the suspect killed the victim with a rifle on Monday afternoon"

A: $((\mathrm{decl\_X}) \Rightarrow \mathrm{story\_X}) \Rightarrow \mathrm{guilty}$

B: $((\mathrm{decl\_Y}) \Rightarrow \mathrm{story\_Y}) \Rightarrow \mathrm{guilty}$

C: $((\mathrm{decl\_X}) \Rightarrow \mathrm{story\_X}) \rightarrow \neg\mathrm{story\_Y}$

D: $((\mathrm{decl\_Y}) \Rightarrow \mathrm{story\_Y}) \rightarrow \neg\mathrm{story\_X}$

Here, C defeats A and D, and D defeats C and B.

Grounded extension: $\emptyset$

Now, do we still want floating conclusions?

Non-admissibility based semantics (1/2)

Why not weaken the requirement of admissibility to, for instance, just conflict-freeness?

For example, why not define an extension as a set *Args* with maximal range ($Args \cup Args^+$)

Advantage: it treats even and odd loops in the same way.

Non-admissibility based semantics (2/2)

Suppose we have the following non-defeasible information: $\{ a, b, \neg(c \wedge d) \}$ as well as two defeasible rules: $a \Rightarrow c$ and $b \Rightarrow d$

Let the strict rules be based on classical entailment

A: $(a) \Rightarrow c$

B: $(b) \Rightarrow d$

C: $((a) \Rightarrow c), \neg(c \wedge d) \rightarrow \neg d$

D: $((b) \Rightarrow d), \neg(c \wedge d) \rightarrow \neg c$

E: $\neg(c \wedge d)$

{A, B, E} is conflict-free,

even though it yields inconsistent conclusions!

References

Argumentation in general:

Prakken & Vreeswijk

Handbook of Philosophical Logic, 2nd edition

"Logics for Defeasible Argumentation"

Reinstatement landmark paper: Dung, AIJ 1995

Quality Postulates: Caminada & Amgoud, AIJ 2007

Argument game for Grounded Semantics

Prakken & Sartor 1997, Caminada Ph.D. thesis

Argument game for Preferred Semantics

Prakken & Vreeswijk, JELIA 2000

"Elements of Argumentation" Besnard & Hunter 2008