Dotting the Is and Crossing the Ts: The 'Arguments' in 'Safety Arguments'

Valentin Cassano and Tom Maibaum

Department of Computing and Software
McMaster University
Canada

Motivation

Certification of Safety Critical Systems

Evidence Based Safety Regimes

Safety Cases

Motivation

A **safety case** is intended to make a compelling case that a system under consideration is adequately safe for its intended purposes through the presentation of an evidence-based **argument**.

The Research Question

When is a safety argument properly formulated?

The Research Question

When is a safety argument properly formulated?

All identified system
hazards have been
mitigated.
Therefore,
the system is safe.

The Research Question

When is a safety argument properly formulated?

All identified system hazards have been mitigated.
Therefore,
the system is safe.

Pigs can fly.
Therefore,
the system is safe.

The Research Question

When is a safety argument properly formulated?

All identified system hazards have been mitigated.
Therefore,
the system is safe.

Pigs can fly.
Therefore,
the system is safe.

The Research Question

When is a safety argument properly formulated?

All identified system hazards have been mitigated.
Therefore,
the system is safe.

Pigs can fly.
Therefore,
the system is safe.

The Research Question

When is a safety argument properly formulated?

All identified system hazards have been mitigated.
Therefore,
the system is safe.

Pigs can fly.
Therefore,
the system is safe.

Our standpoint is that the question above **must** be **approached** from an **inferential** point of view.

Outline of the Presentation

Arguments

Safety Arguments

Discussion

Conclusions and Further Work

Arguments

The **ordinary** understanding of an argument

> An **argument** is a set of assertions in which one or more of them, the **premisses**, are put forward so as to offer a rationale for another assertion, the **conclusion**.

Arguments

The **ordinary** understanding of an argument

An **argument** is a set of assertions in which one or more of them, the **premisses**, are put forward so as to offer reasons for another, the **conclusion**.

The trip system is triggered correctly and in a timely fashion if the temperature of the reactor reaches a critical level. This is something that has been established under adequate test conditions. The previous is justified as follows: if the temperature of the reactor were to reach a critical level with the trip system not being triggered, then, the test conditions would be proven inadequate, for such a problem should have been discovered during testing. But this is a contradiction for the test conditions indeed are adequate.

Arguments

The **ordinary** understanding of an argument

An **argument** is a set of assertions in which one or more of them, the **premisses**, are put forward in support of, or offered reasons for another assertion, the **conclusion**.

The trip system is triggered correctly and in a timely fashion if the temperature of the reactor reaches a critical level. This is something that has been established under adequate test conditions. The previous is justified as follows: if the temperature of the reactor were to reach a critical level with the trip system not being triggered, then, the test conditions would be neither adequate for such a problem should have been discovered during testing. But this is a contradiction for the test conditions to be adequate.

The **inferential** understanding of an argument

An **argument** is a series of assertions in which: every assertion other than the last one, the **conclusion**, is either assumed to be a **premiss** and taken as being the case, or obtained from some foregoing assertions in this series by virtue of a **rule of inference**.

Arguments

The **ordinary** understanding of an argument

An **argument** is a set of assertions in which one of them, the **premisses**, are put forward so as to offer reasons for another assertion, the **conclusion**.

The **inferential** understanding of an argument

An argument is a series of assertions in which: every assertion other than the last one, the **conclusion**, is either assumed to be a **premiss**, and taken as being the case, or obtained from some foregoing assertions in this series by virtue of a **rule of inference**.

Inference Systems

Classes of arguments formulate **inference systems**, a.k.a., **logics**.

The chief aim of a logic **L** is to say, accurately and systematically, what **being a consequence of** or **following from** amounts to by **stating precisely** which **arguments** are **properly formulated**.

The trip system is triggered correctly and in a timely fashion if the temperature of the reactor reaches a critical level. This is something that has been established under adequate test conditions. The previous is justified as follows: if the temperature of the reactor were to reach a critical level with the trip system not being triggered, then, the test conditions would be proven inadequate, for such a problem should have been discovered during testing. But this is a contradiction for the test conditions indeed are adequate.

Arguments

The **ordinary** understanding of an argument

An **argument** is a set of assertions in which one or more of them, the **premisses**, are put forward so as to offer reasons for another assertion, the **conclusion**.

The **inferential** understanding of an argument

An argument is a series of assertions in which: every assertion other than the last one, the **conclusion**, is either assumed to be a **premiss** and taken as being the case, or obtained from some foregoing assertions in this series by virtue of a **rule of inference**.

Inference Systems

Classes of arguments formulate **inference systems**, a.k.a., **logics**.

The chief aim of a logic **L** is to say, accurately and systematically, what **being a consequence of** or **following from** amounts to by **stating precisely** which **arguments** are **properly formulated**.

The trip system is triggered correctly and in a timely fashion if the temperature of the reactor reaches a critical level. This is something that has been established under adequate test conditions. The previous is justified as follows: if the temperature of the reactor were to reach a critical level with the trip system not being triggered, then, the test conditions would be proven inadequate, for such a problem should have been discovered during testing. But this is a contradiction for the test conditions indeed are adequate.

Arguments

The **ordinary** understanding of an argument

An **argument** is a set of assertions in which one or more of them, the **premisses**, are put forward so as to offer reasons for another assertion, the **conclusion**.

The **inferential** understanding of an argument

An argument is a series of assertions in which: every assertion other than the last one, the **conclusion**, is either assumed to be a **premiss** and taken as being the case, or obtained from some foregoing assertions in this series by virtue of a **rule of inference**.

Inference Systems

Classes of arguments formulate **inference systems**, a.k.a., **logics**.

The chief aim of a logic **L** is to say, accurately and systematically, what **being a consequence of** or **following from** amounts to by **stating precisely** which **arguments** are **properly formulated**.

The trip system is triggered correctly and in a timely fashion if the temperature of the reactor reaches a critical level. This is something that has been established under adequate test conditions. The previous is justified as follows: if the temperature of the reactor were to reach a critical level with the trip system not being triggered, then, the test conditions would be proven inadequate, for such a problem should have been discovered during testing. But this is a contradiction for the test conditions indeed are adequate.

The argument above is potentially **fallacious** for it may misuse the rule of **reductio ad absurdum**.

Such a judgement call is made by analyzing its structure and coming to the conclusion that said rule of inference does not accommodate for its formulation unless the test conditions are established to be adequate by some other means.

Safety Arguments

The **elements** of a **safety case**

By no means exhaustive, a sensibly good safety case is, at a bare minimum, comprised of the following elements:

1. A well-defined **safety life-cycle**.
2. A rigorous **hazard analysis**.
3. Adequate **safety goals**.
4. The production and collection of **safety related evidence**.
5. The structuring of evidence and safety goals in the form of a **safety argument**.

Safety Arguments

The **elements** of a **safety case**

By no means exhaustive, a sensibly good safety case is, at a bare minimum, comprised of the following elements:

1. A well-defined **safety life-cycle**.
2. A rigorous **hazard analysis**.
3. Adequate **safety goals**.
4. The production and collection of **safety related evidence**.
5. The structuring of evidence and safety goals in the form of a **safety argument**.

Of all the elements of a **safety case** that are worth paying attention to, here in particular, we pay close attention to the notion of a **safety argument**

Safety Arguments

The **elements** of a **safety case**

By no means exhaustive, a sensibly good safety case is, at a bare minimum, comprised of the following elements:

1. A well-defined **safety life-cycle**.
2. A rigorous **hazard analysis**.
3. Adequate **safety goals**.
4. The production and collection of **safety related evidence**.
5. The structuring of evidence and safety goals in the form of a **safety argument**.

Of all the elements of a **safety case** that are worth paying attention to, here in particular, we pay close attention to the notion of a **safety argument**

And more specifically, to their formulation within the scope of an **inference system** for **safety argumentation**

Safety Arguments

The **elements** of a **safety case**

By no means exhaustive, a sensibly good safety case is, at a bare minimum, comprised of the following elements:

1. A well-defined **safety life-cycle**.
2. A rigorous **hazard analysis**.
3. Adequate **safety goals**.
4. The production and collection of **safety related evidence**.
5. The structuring of evidence and safety goals in the form of a **safety argument**.

Of all the elements of a **safety case** that are worth paying attention to, here in particular, we pay close attention to the notion of a **safety argument**

And more specifically, to their formulation within the scope of an **inference system** for **safety argumentation**

This is our sought after ideal

Safety Arguments

Comment on the shortcomings of classical logic as a logic for safety argumentation, uncertainty, defeasibility, etc.

Classical Logic is **inadequate**

Given evidence **e**, presumably, **c** is the case. Since E, who is considered an expert in the domain in which **c** occurs, has claimed that based on **e**, **c** is the case. On account of E having presented some credentials attesting to his expertise. Unless, E's credentials are inadequate or **e** is vitiated.

The next slide (still missing) should provide a brief introduction as to how Toulmin can help to cope with the above perceived deficiency.

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

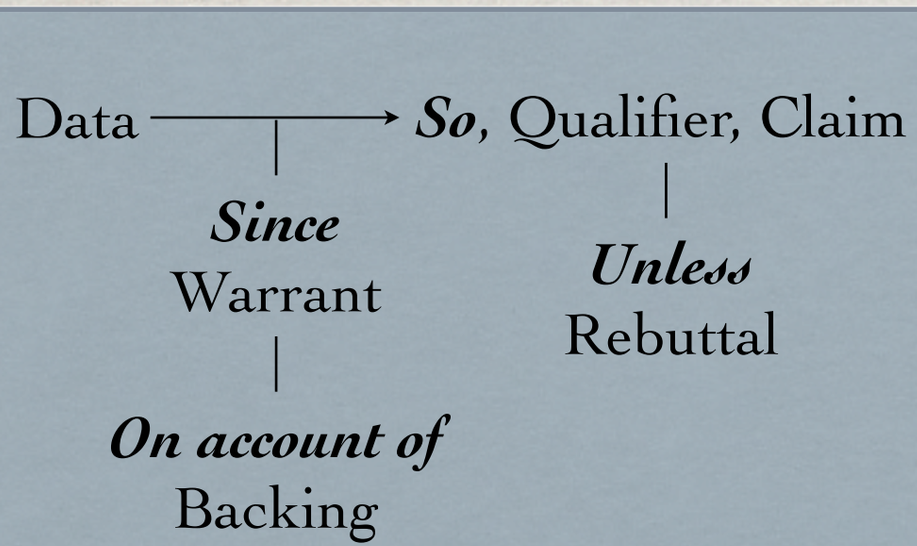
Comment on the shortcomings of classical logic as a logic for safety argumentation, uncertainty, defeasibility, etc.

The next slide (still missing) should provide a brief introduction as to how Toulmin can help to cope with the above perceived deficiency.

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

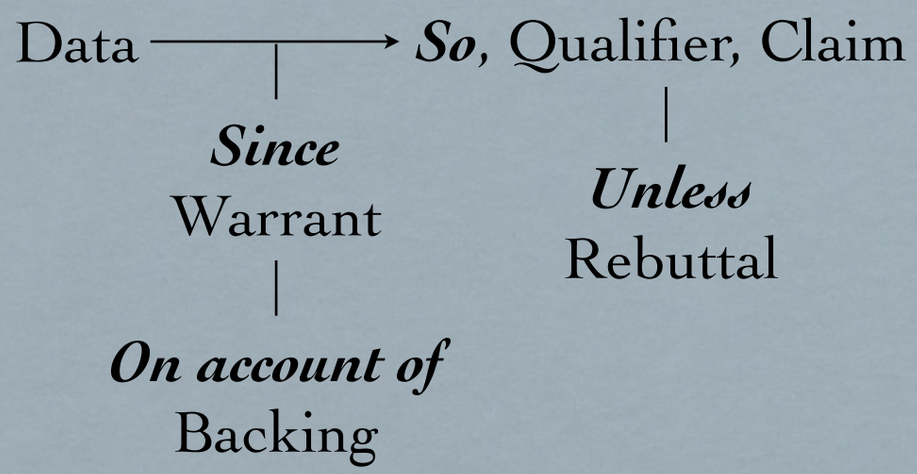
Toulmin's **argument pattern**



Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**



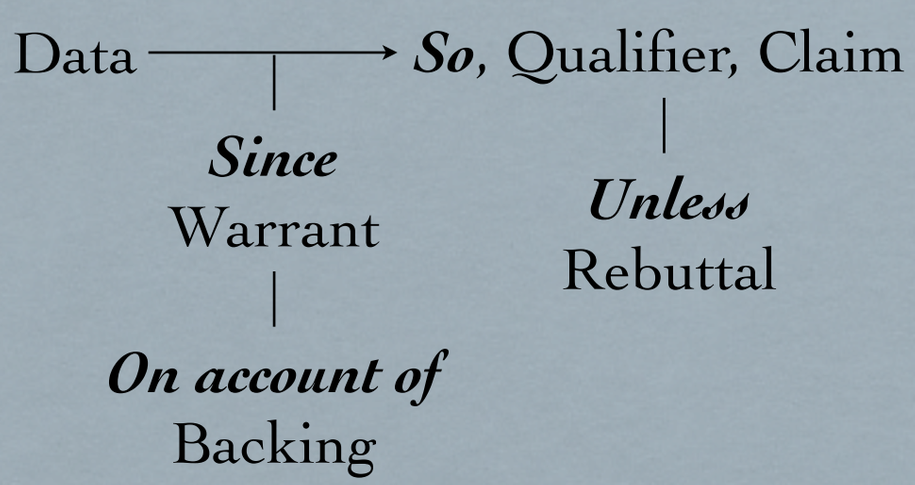
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**



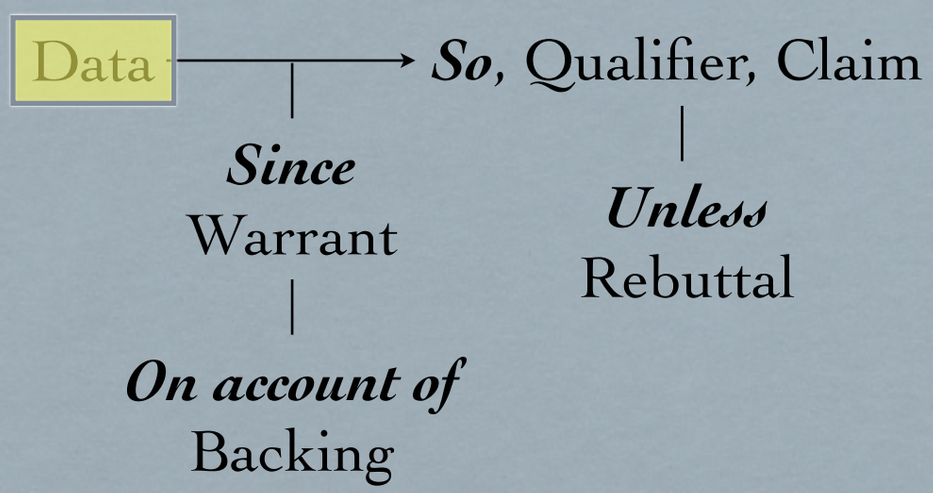
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**



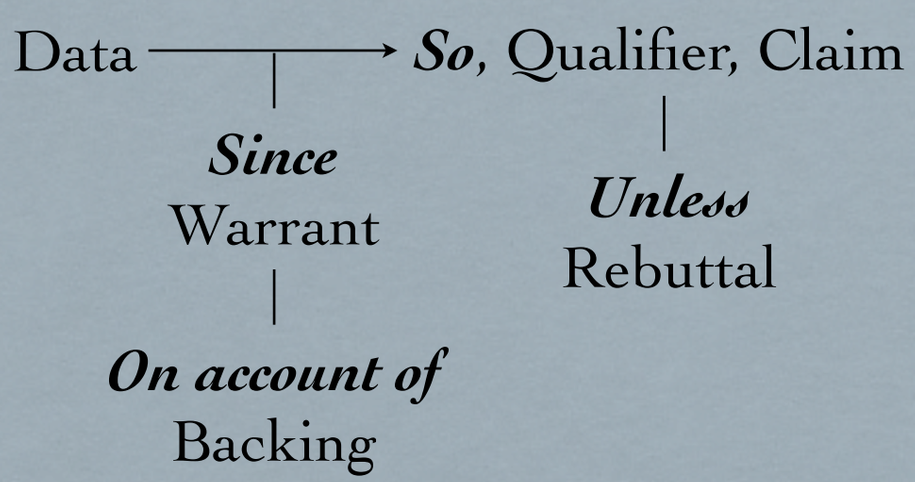
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**



Given evidence **e**

Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

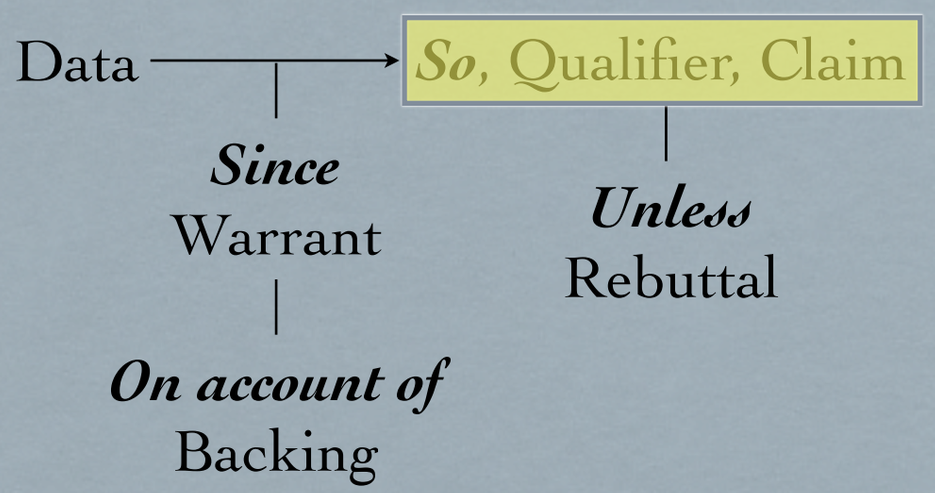
Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Given evidence **e**

Toulmin's **argument pattern**



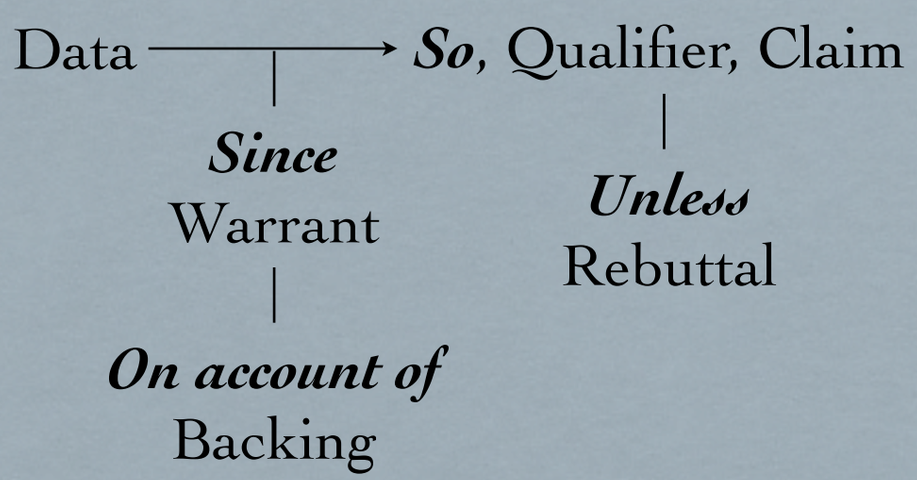
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**



Given evidence **e** ⟶ ***So***, presumably, **c**

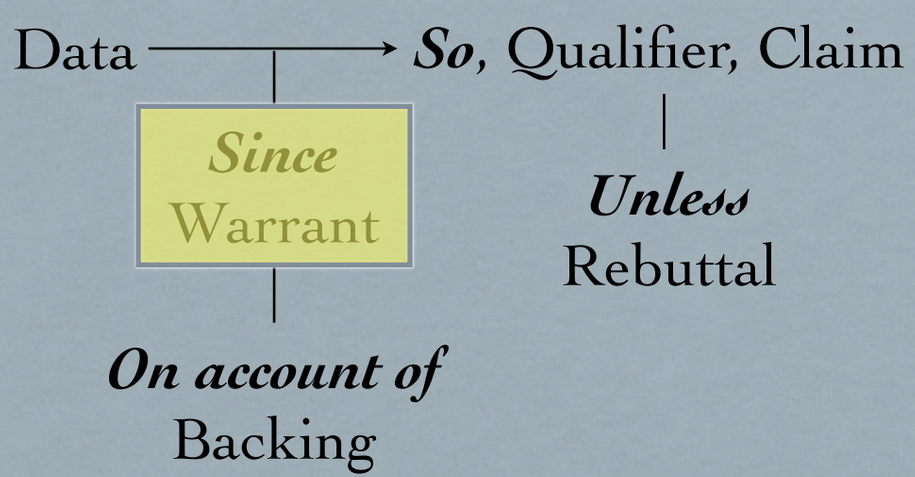
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since E, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of E having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**



Given evidence **e** ⟶ ***So***, presumably, **c**

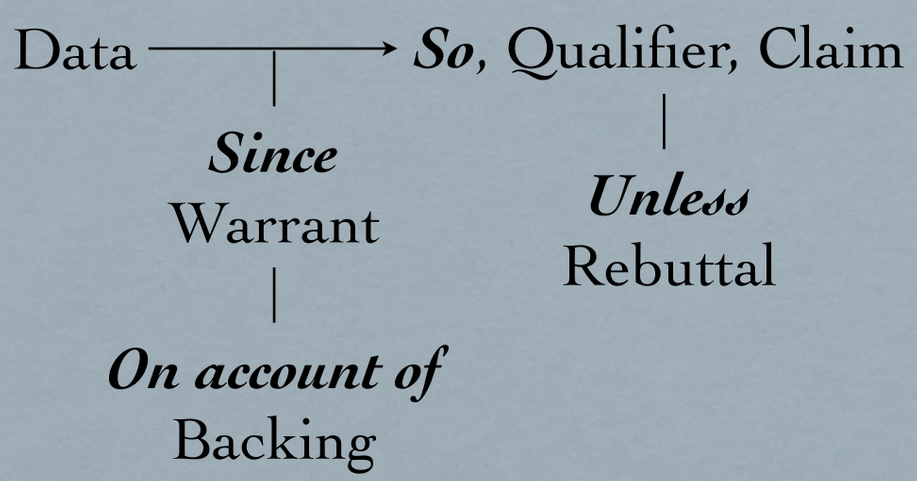
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

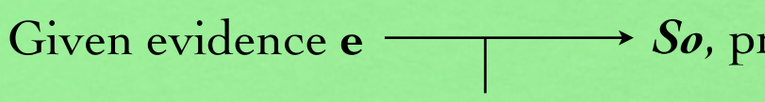
Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**





Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

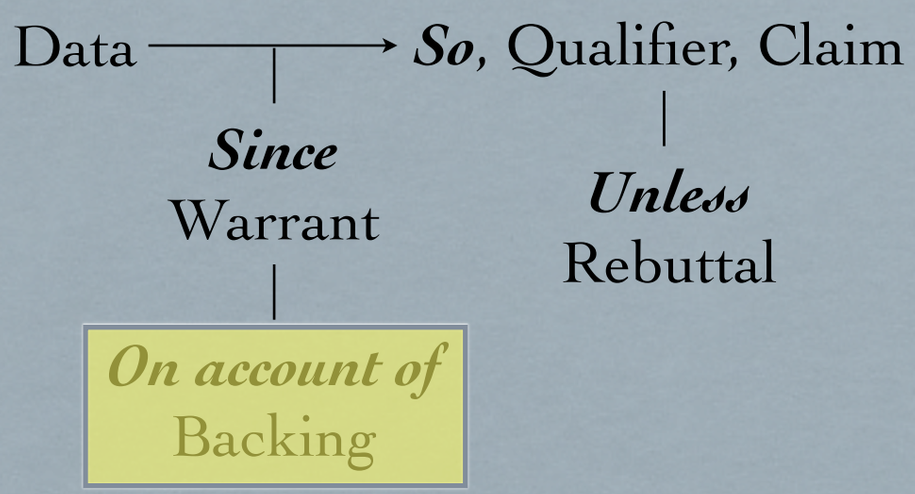
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

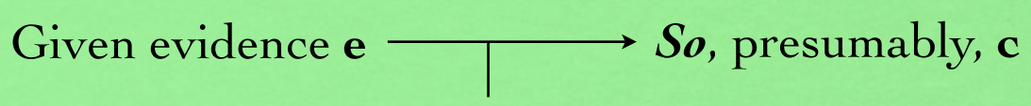
Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of E having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**





Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

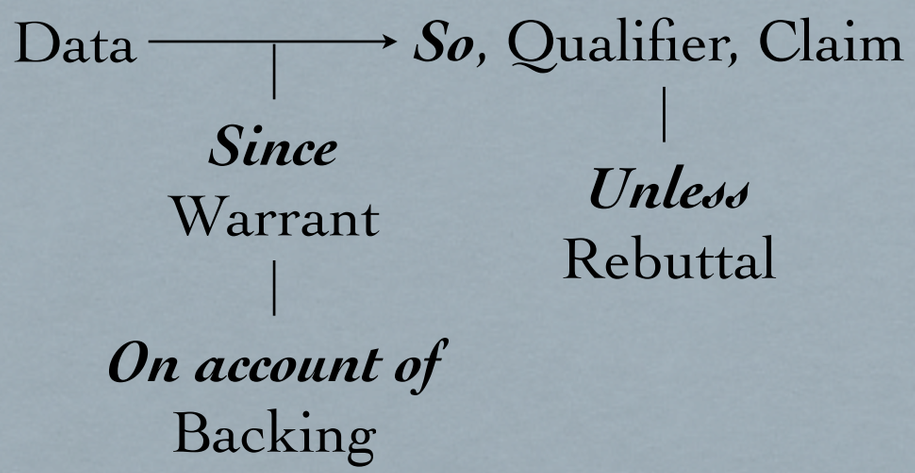
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

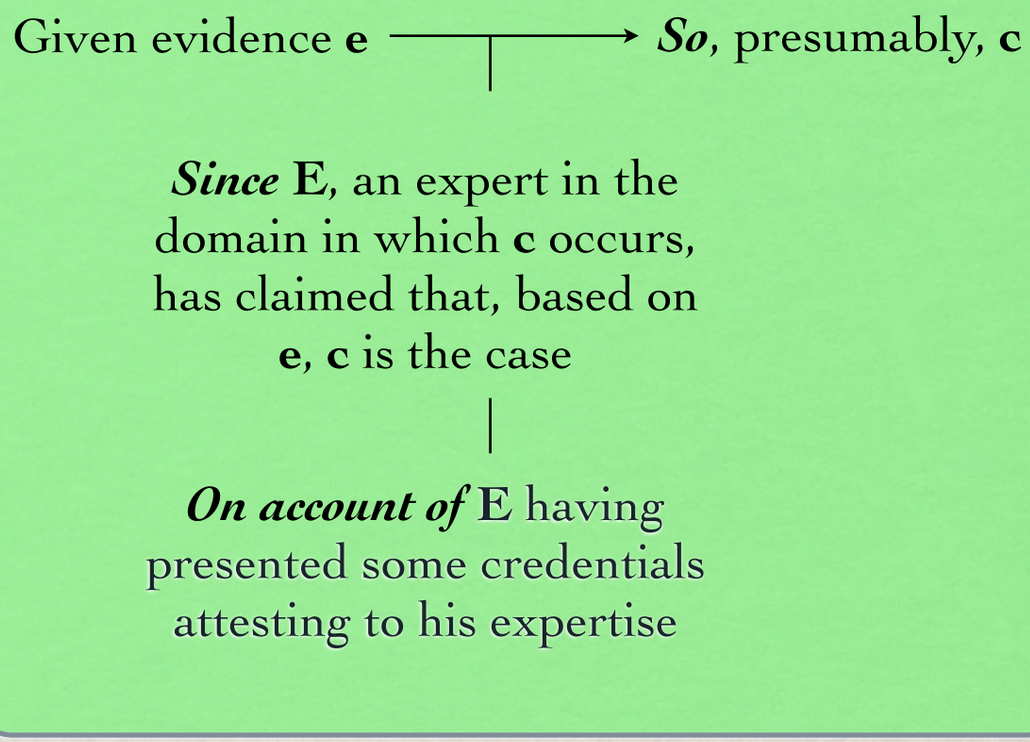
Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**





Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

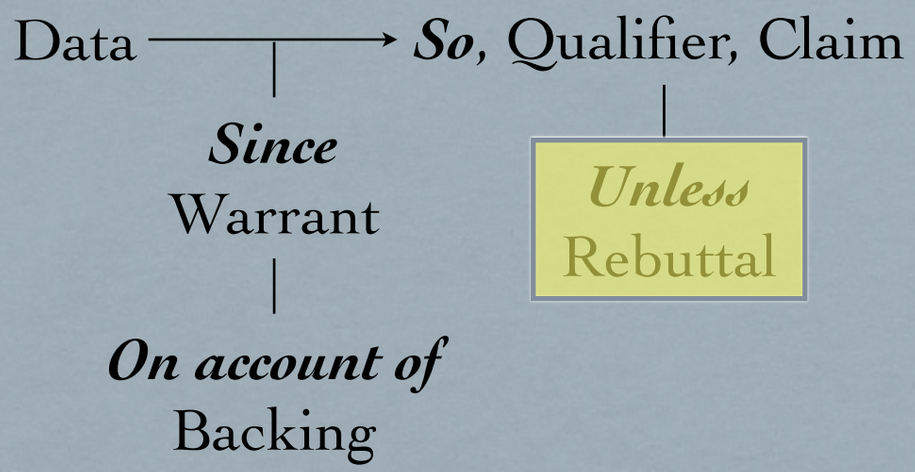
Discover **fallacies** in safety arguments

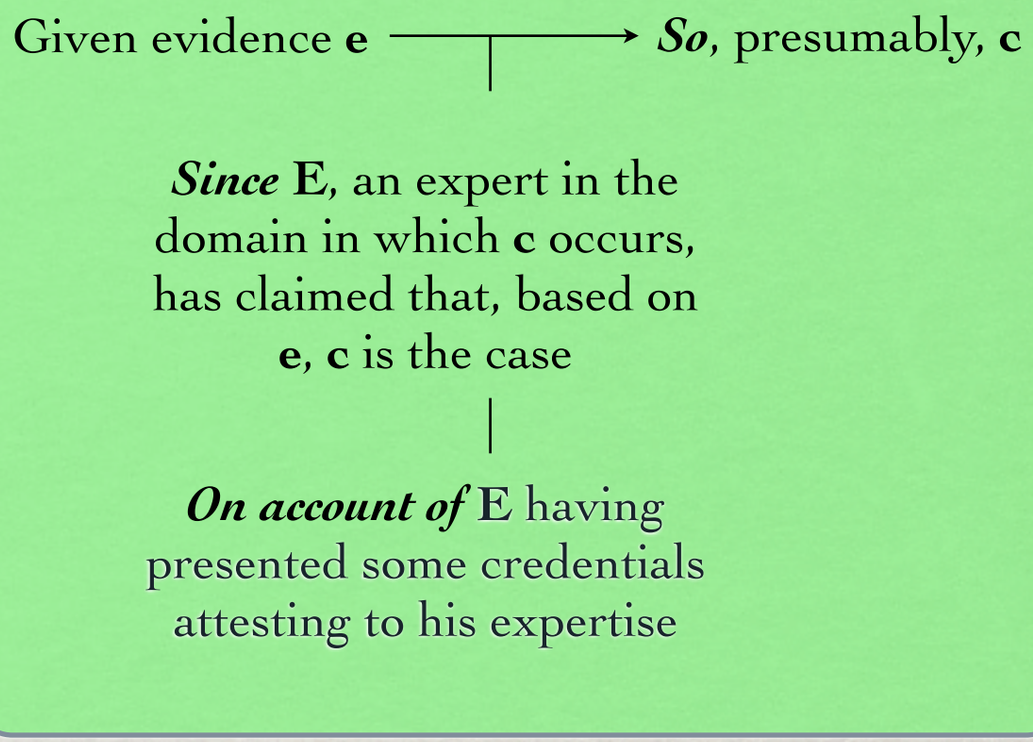
Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.

Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**





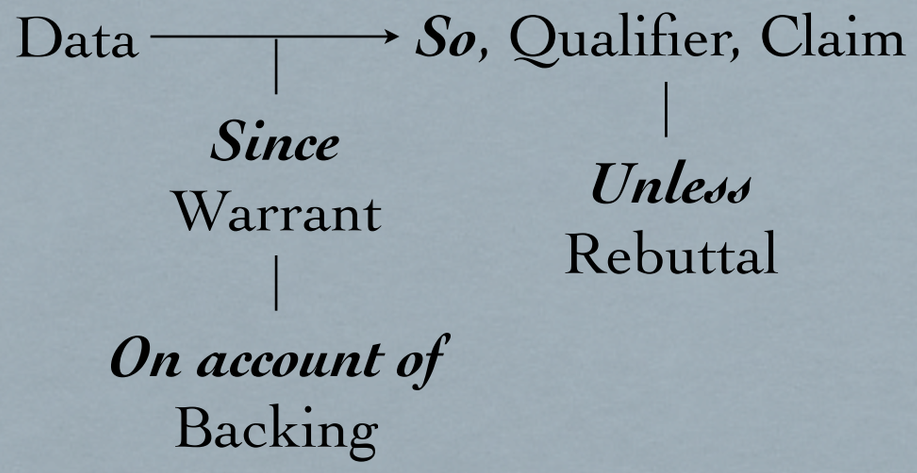
Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation**.

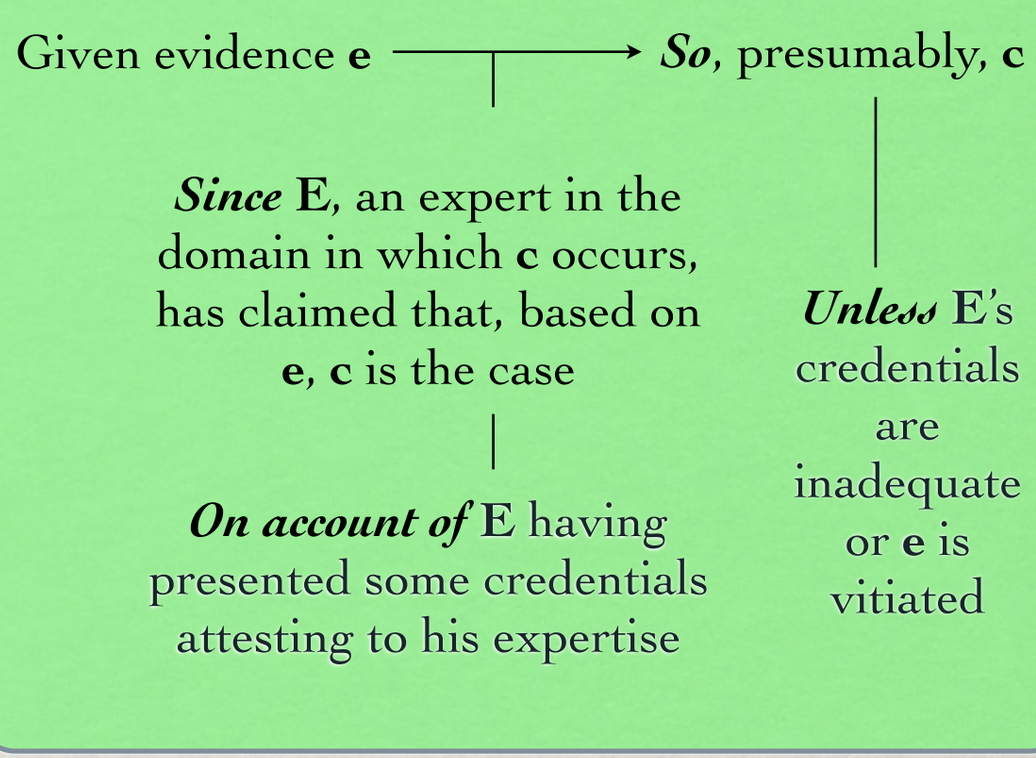
Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e**,
presumably, **c** is the case.
Since **E**, who is considered an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case.
On account of **E** having presented some credentials attesting to his expertise.
Unless,
E's credentials are inadequate or **e** is vitiated.

Toulmin's **argument pattern**





Toulmin's notion of an argument pattern establish a scientific and rigorous **basis** on which to construct an **inference system** for **safety argumentation.**

Discover **fallacies** in safety arguments

Safety Arguments

Given evidence **e** → *So*, presumably, **c**

Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

On account of **E** having presented some credentials attesting to his expertise

Unless **E**'s credentials are inadequate or **e** is vitiated

Safety Arguments

Given evidence **e**

is **e** sufficient?

Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

On account of E having presented some credentials attesting to his expertise

So, presumably, **c**

Unless E's credentials are inadequate or **e** is vitiated

Safety Arguments

Given evidence **e**

is **e** sufficient?

Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

has **E** usually been right in his claims?

On account of **E** having presented some credentials attesting to his expertise

So, presumably, **c**

Unless **E**'s credentials are inadequate or **e** is vitiated

Safety Arguments

Given evidence **e**

is **e** sufficient?

Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

has **E** usually been right in his claims?

On account of **E** having presented some credentials attesting to his expertise

is **E** a member of a recognized authority?

So, presumably, **c**

Unless **E**'s credentials are inadequate or **e** is vitiated

Safety Arguments

Given evidence **e**

is **e** sufficient?

Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

has **E** usually been right in his claims?

On account of **E** having presented some credentials attesting to his expertise

is **E** a member of a recognized authority?

how are **E**'s credentials evaluated

So, presumably, **c**

Unless **E**'s credentials are inadequate or **e** is vitiated

Safety Arguments

Given evidence **e**

is **e** sufficient?

Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

has **E** usually been right in his claims?

On account of **E** having presented some credentials attesting to his expertise

is **E** a member of a recognized authority?

how are **E**'s credentials evaluated

So, presumably, **c**

Unless **E**'s credentials are inadequate or **e** is vitiated

how would these affect the validity of **c**?

Safety Arguments

what is **c**'s confidence value?

Given evidence **e** → *So*, presumably, **c**

is **e** sufficient?

Since **E**, an expert in the domain in which **c** occurs, has claimed that, based on **e**, **c** is the case

Unless **E**'s credentials are inadequate or **e** is vitiated

has **E** usually been right in his claims?

is **E** a member of a recognized authority?

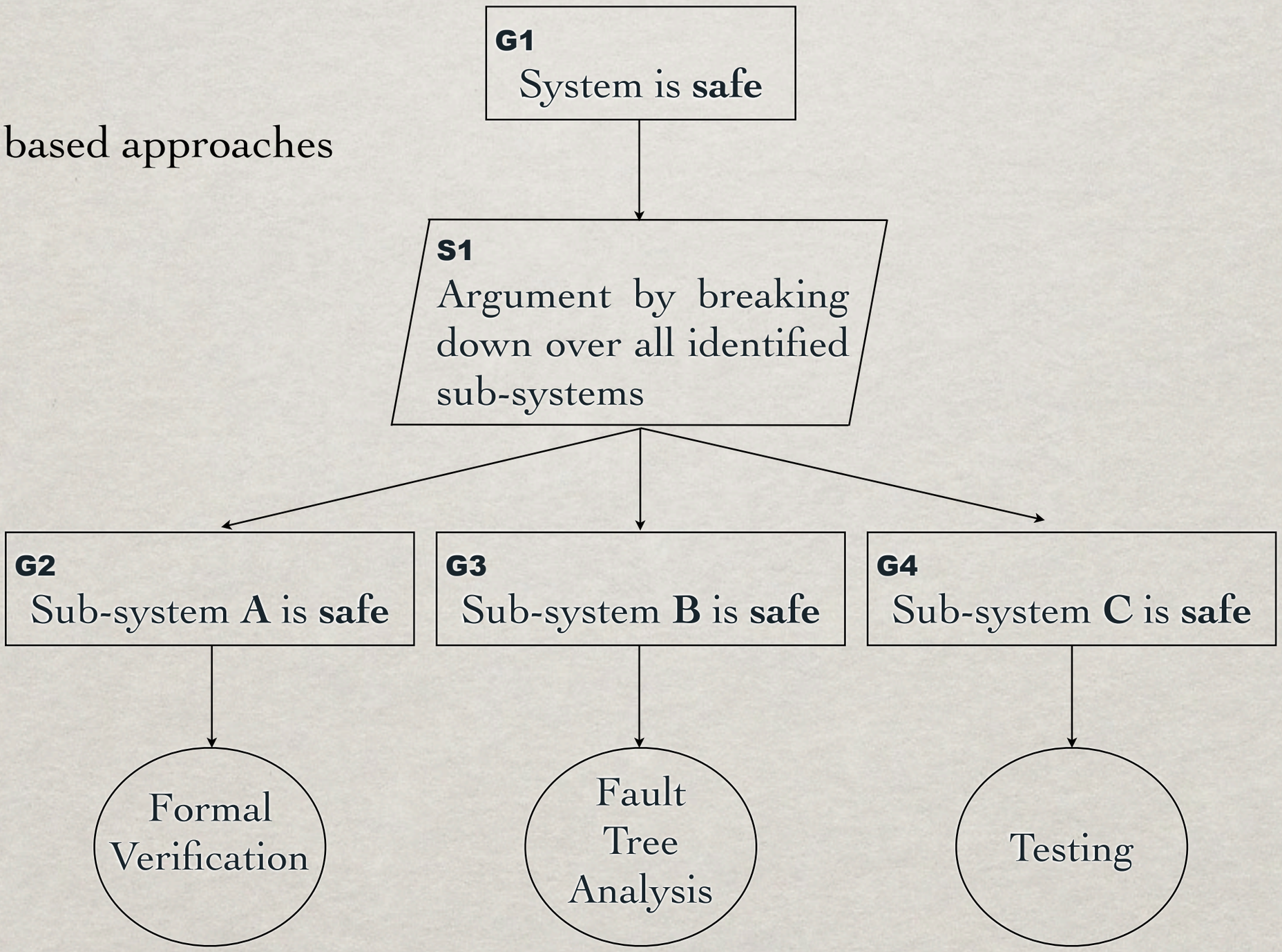
On account of **E** having presented some credentials attesting to his expertise

how would these affect the validity of **c**?

how are **E**'s credentials evaluated

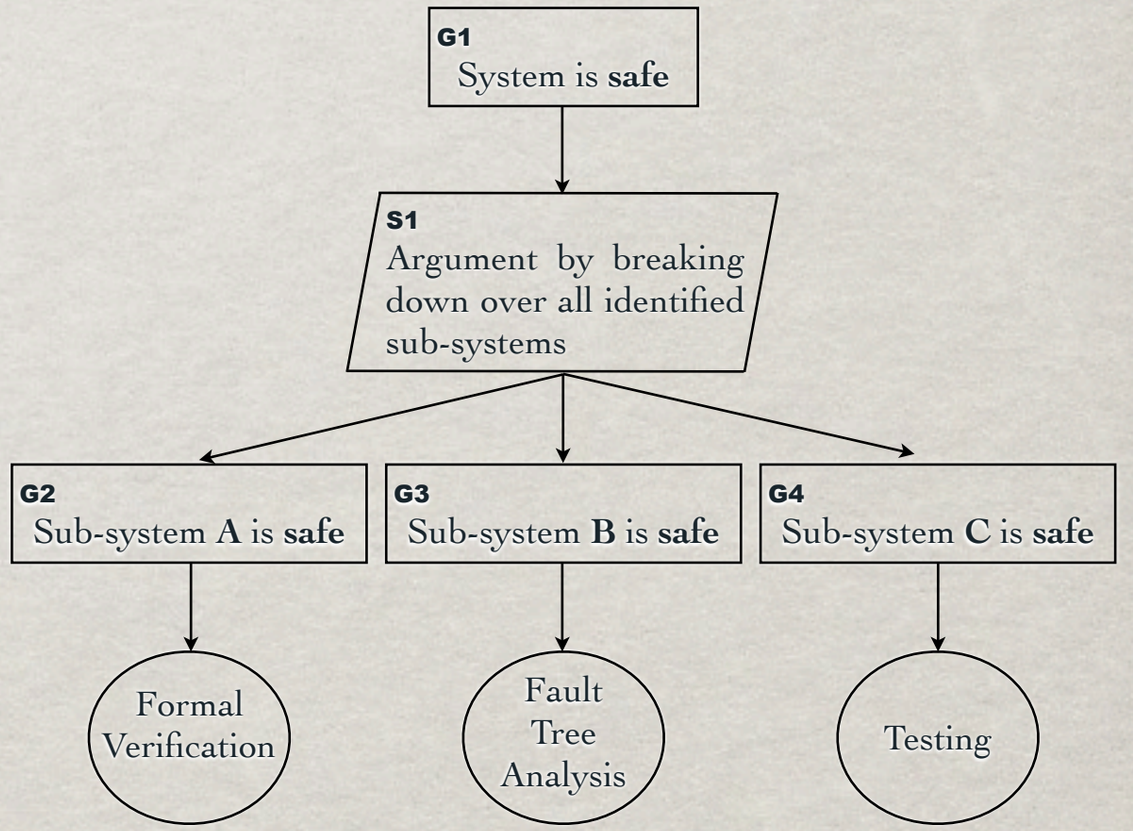
Discussion

GSN based approaches



Discussion

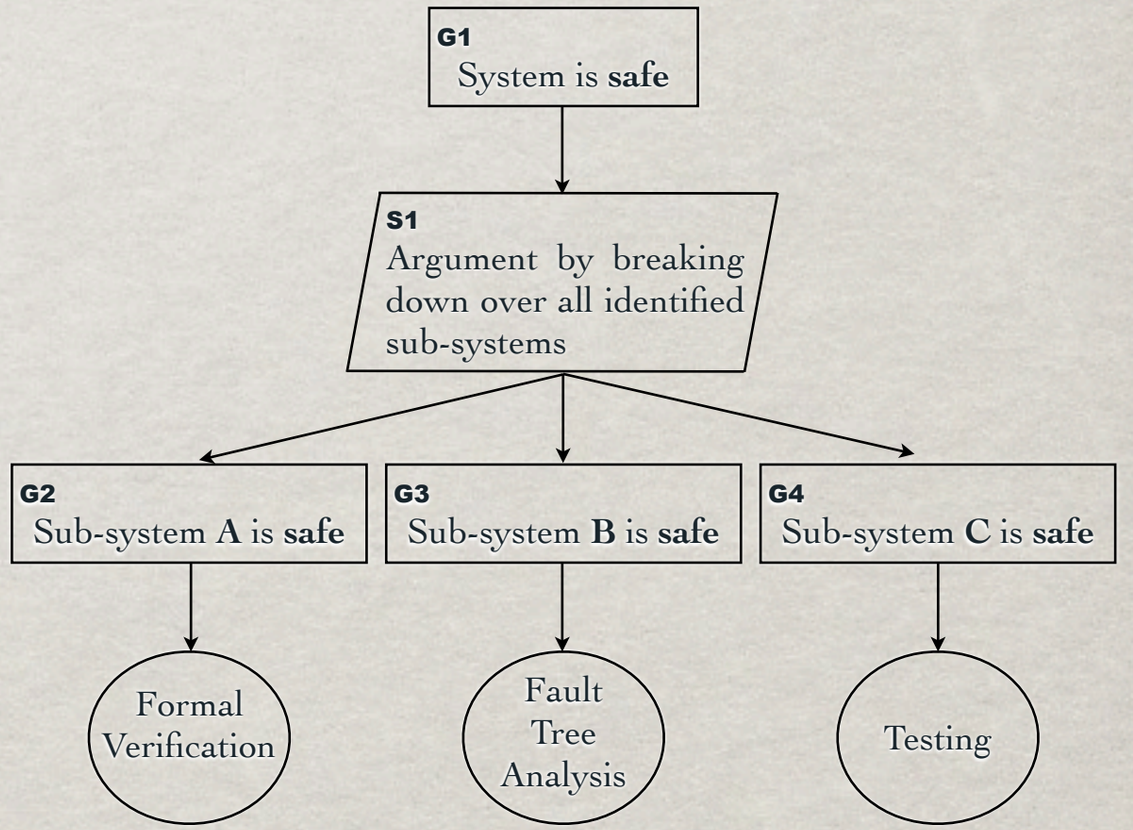
GSN based approaches



Discussion

GSN based approaches

In essence, safety goal decomposition is a **breaking down** approach to the **design** of a safe system (pretty much analogous to problem solving by decomposition).

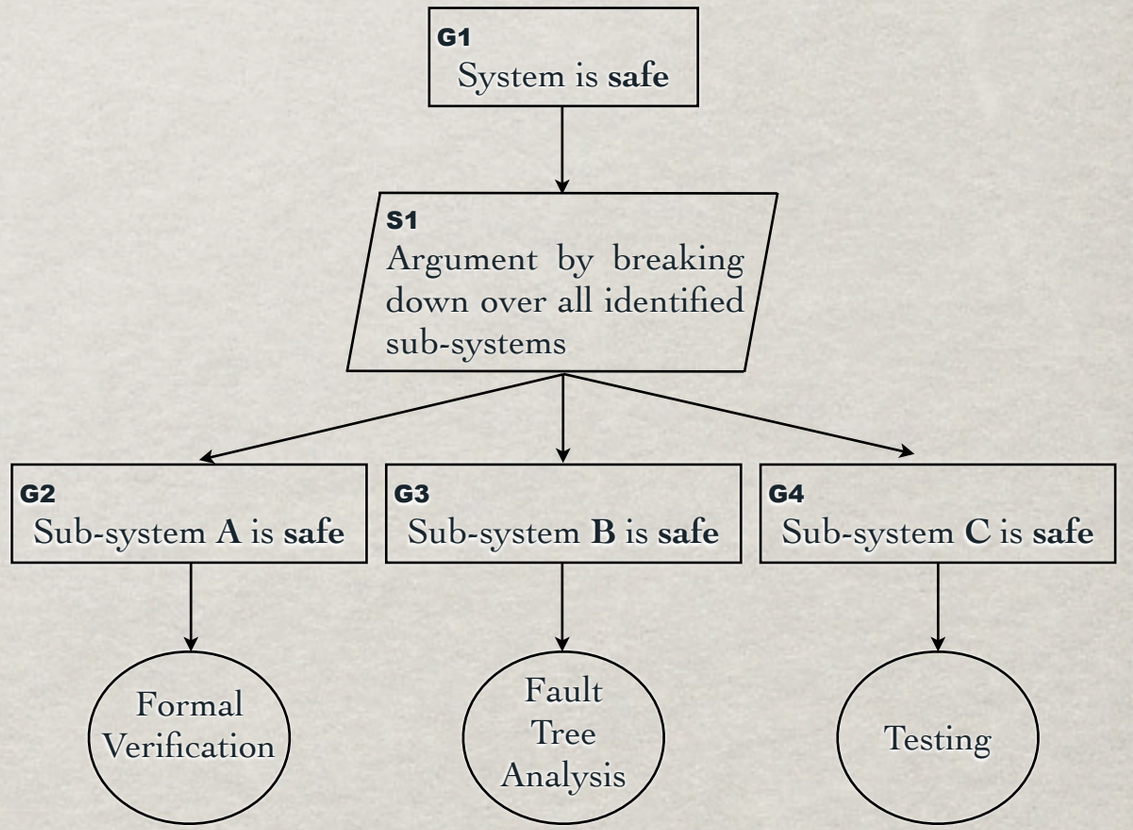


Discussion

GSN based approaches

In essence, safety goal decomposition is a **breaking down** approach to the **design** of a safe system (pretty much analogous to problem solving by decomposition).

Instead, safety argument formulation is a **ground up** approach that moves from some produced evidence to a safety claim under consideration to establish that such a safety claim is fulfilled.



Summary and Further Work

When is a safety argument properly formulated?

All identified system hazards have been mitigated.
Therefore,
the system is safe.

Pigs can fly.
Therefore,
the system is safe.

Our standpoint is that the question above **must** be **approached** from an **inferential** point of view.

Arguments

Safety Arguments

Discussion

The actual formulation of an

Inference System

for

Safety Argumentation

Questions

Tom Lehrer - That Was the Year That Was (1965)

Any ideas expressed on this record should not be taken as representing Mr. Lehrer's

(or for the purposes of this presentation Dr. Maibaum's)

true convictions, for indeed he has none. "**If anyone objects to any statement I make," he has said, "I am quite prepared not only to retract it, but also to deny under oath that I ever made it."**