2.38. Fundamentals of Indirect Deduction

1. Indirect Deduction: The Philosophy. Recall that we earlier recast our definition of "validity counterexample" in terms of an argument's **counterexample set**.[1]

> **Counterexample Set** for an argument: the set of sentences
> {Premises, Negation of Conclusion}

A validity counterexample for the argument – a possible situation making all the premises true and the conclusion false – will make every sentence in the counterexample set true. In that case, the counterexample set is **consistent** (simultaneously satisfiable). And that provided an alternate definition of "invalid argument".

> An **invalid argument** is an argument whose counterexample set is consistent.

A valid argument, by contrast, is one lacking any validity counterexample – meaning that its Counterexample Set is unsatisfiable, hence **inconsistent**.

> A **valid argument** is an argument whose counterexample set is inconsistent.

Now in exploring the links between validity and inconsistency, we noted moreover that an inconsistent set of sentences logically entails contradictory sentences: **some sentence, and also its negation**. Since a consistent set of sentences cannot entail such an inconsistent pair of sentences, this entailment feature is the hallmark of an inconsistent set of sentences.[2]

> A set of sentences is **inconsistent** if (and only if) it logically entails opposites: some sentence, and its negation.

Putting these observations together, we see that **an argument is valid if (and only if) its counterexample set entails some sentence and also its negation**.

---

[1] In 2.20.

[2] In 2.18.1, Problem G.

Applying this approach to deduction yields an **indirect deduction**.

> **Fundamental Principle of Indirect Deduction**
>
> **An argument is valid if (and only if) its counterexample set {Premises, Negation of Conclusion} entails some sentence, and also that sentence's negation.**

An example of indirect deduction comes from a familiar argument whose validity has been demonstrated several times over.

1. (P ∨ Q)
2. ~Q

∴ P

Now we can show that this argument is valid by deducing opposite sentences from the premises combined with the negation of the conclusion.

1. (P ∨ Q)	Premise
2. ~Q	Premise
3. ~P	Negation of conclusion

∴ (Some sentence, and its negation)

That's simple enough: Sentences (1) and (3) yield "Q" by ∨–; and "Q," along with the second premise "~Q," are opposites.

1. (P ∨ Q)	Premise
2. ~**Q**	Premise
3. ~P	Negation of conclusion
4. **Q**	1, 3, ∨**–**

It turns out that indirect deductions can do anything ordinary deductions can do – **and more**. The following, for instance, looks like a model example of a valid argument.[3]

1. We’re not having both ice cream and cake.	1. $\sim(P \wedge Q)$
2. We’re having ice cream.	2. P
∴ We’re not having cake.	∴ $\sim Q$

But good luck deducing that conclusion from the premises, using just our seven rules: we’re out of Elim rules from the beginning, and Intro rules won’t yield “~Q”. By contrast, it’s quite simple to show indirectly that this argument is valid.

2. Indirect Deduction: Technical Techniques. What remains is just bookkeeping details.

A deduction has so far begun with premises, followed by a “Get” line for the conclusion. That won’t change with indirect deductions. But now we add another sentence, the negation of that conclusion. This isn’t an additional premise, but rather an **assumption** entertained solely to demonstrate how noxious the consequences would be, *were* we to accept it.

It’s rather like a situation where, faced with a fork in the road, we appeal to a roadmap to consider, hypothetically, what **would** happen **if** we took the left road. If the roadmap reports that a left turn leads off an unfinished bridge to certain death, we don’t conclude that we actually have driven off the bridge and died – only that this is what **would** happen, if we **were** to do so. And precisely because of these unacceptable side-effects, we reject such a route.

Likewise with an indirect deduction: the negation of the conclusion isn’t a sentence actually accepted, the way a premise is – merely one entertained hypothetically, to see where it leads.

We call this the **Assumption of an Indirect Deduction –** or “**AID**”, for short. We mark this assumption, and its various consequences, as **purely**

[3] Indeed, it was treated as a basic form of inference by the ancient Stoic logicians; see (Mates 19XX: xxx).

hypothetical by placing them in a **box**, as follows. (As an added reminder that we're constructing an indirect deduction, we can write "(ID)" on the "Get" line.)

1. Premise 1
2. Premise 2

Get: Conclusion (ID)

3. Negation of Conclusion AID

Supposing that premises, AID, and deductive rules lead to a contradictory pair of sentences, our hypothetical reasoning draws to an end: since inconsistency in logic is even more unacceptable than a trip off a bridge in the motoring world, we reject the AID on account of its noxious consequences. But the assumption has served its purpose: since premises plus negation of conclusion have now been shown to yield contradictory sentences, the original argument must be **valid**. So we end the deduction by asserting the conclusion.

1. Premise 1
2. Premise 2

~~Get~~: Conclusion (ID)

3. Negation of Conclusion AID

.

.

.

9. Sentence

10. Negation of That Sentence

11. Conclusion 3, 9, 10 ID

That's why we justify the conclusion the way we do on the final line: we cite the **original assumption** (the AID – here, line 3), and the two contradictory sentences which that assumption led to (here lines 9 and 10).

And just as we don’t conclude from hypothetical musings over a map that we’ve actually driven off a bridge, so with indirect deductions we don’t believe the contradictory consequences deduced from the AID, within the ID box. Once the hypothetical reasoning is complete, and the ID box is closed, all sentences within that box are lost to us. Technically, such sentences are **unusable** – meaning that **no rules of inference can be applied to them**.

Sentences in a closed box are **unusable**.

But surrendering such hypotheticals is worth the trade, since by closing the ID box we achieve our ultimate goal: the conclusion of the argument.

As a final notational detail, we relax how we draw ID boxes. While the examples surveyed so far have all fit comfortably within the boxes we’ve drawn, we may start with an ID box that ends up too narrow to fit some later sentence.

1. Premise 1
2. Premise 2

Get: Conclusion (ID)

3. | Negation of Conclusion | AID
4. | This line fits. |
5. | But this line is too long to fit in the box.

So we’ll leave off the right side of ID boxes, in anticipation of long later lines.

1. Premise 1
2. Premise 2

Get: Conclusion (ID)

3. | Negation of Conclusion AID
4. | This line fits.
5. | Now even this very long line fits.

We will, however, still call this more open structure a "box". And we'll continue to close such a 'box' with a horizontal line across the bottom.

1.	Premise	
		~~Get~~: Conclusion (ID)
2.	Assumption	AID
	.	
	.	
	.	
7.	Some Sentence	
8.	Negation of That Sentence	
9.	Conclusion	2, 7, 8, ID

And as before, lines in a **closed** 'box' are unusable (except to justify that ID – as on Line 9, above).[4]

> "All these things, O God, are conceived with forethought, born with determination, nursed with exactness, governed by rules, directed by reason, and then slain and buried after a prescribed method. And even their silent graves that lie within the human soul are marked and numbered."
>
> Kahlil Gibran, "The Perfect World," in **The Madman**

[4] Question: if the sentences inside a closed ID box are unusable, why are we allowed to cite them in our justification of the conclusion (below the ID box)? Answer: Calling sentences unusable means only that **rules of inference** can't apply to them. However **ID isn't a rule of inference**, but rather a special type of **deduction**.

Summary: Indirect Deduction (ID)

- Write (**ID**) next to the “Get” line, as a reminder.
- Immediately following the “Get” line, begin a **box**, in which the Indirect Deduction occurs.
- The first line in the ID box is the **Assumption of the Indirect Deduction** (**AID**): the negation of the sentence on the “Get” line.
- Using deductive rules on all available lines (premises and AID), deduce (i) some sentence and (ii) its negations.
- Once these two sentences have been deduced, **close** the ID box.
- Note: once an ID box is closed, no rules of inference can be applied to any line in that box. **Sentences inside a *closed* ID box are unusable**.
- Beneath the ID box write the conclusion of the argument (the sentence on the “Get” line). The justification for this conclusion cites three lines: the **AID**, and the two lines constituting **some sentence, and its negation.** These three line numbers are followed by “**ID**”.