Paradigm Academy Brian Burns & James Dama

Chart Logic & Core Mechanics

August 2018

Introduction

This document explains the logic and core mechanics of CT charts. After a charting session you may have seen a complicated diagram with many colored boxes and arrows that purportedly describes part of a human mind. Using the language of CT, this document explains what each one of the boxes and arrows are, what the logical relations are between them, and what one should practically do when navigating a chart during a charting session.

In contrast to what this document is, it may be useful to note what this document is not. This document is not an explanation of CT: for that, read *Connection Theory: Theory and Practice* and *CT: a Crash Course*. Indeed, many of the basic concepts such as the belief rule and the IG constraint will be assumed as known in this document. This is also not an original research document. It is rather a technical reference text for a years-old practice with many contributors. Finally, it is also not a charting how-to manual. While understanding the contents of this document is necessary for charting proficiency, it is not sufficient. For that one must gain intimate knowledge of mental structures not codified in the aforementioned documents such as layer structures and IPs, alongside implicit models built from charting actual people.

It may also be useful to note why reading this document is a useful thing to do. First, it teaches you what a chart is: one cannot progress as a trainer without understanding the mechanics of charting. Second – and with emphasis – clearly and precisely understanding chart logic and notation will immensely help your debugging practice by providing a paradigm through which to interpret and pattern match introspective data. After reading this document, we hope that you begin to see the relations that correspond to grey chains, attainment responses, the arrow necessity condition, and many other later-described items when you introspect, even without the notational aid of a chart.

With this in mind, we hope the reader is motivated to move forward. In chapter one, we encounter a full chart and enumerate its components to be later studied. In chapter two, we examine the bulk constituents of a chart – grey and white chains – which track logically-related sequences of actions that are respectively evaluated as bad or good to do. In chapter three, we enumerate constrained belief structures that are consequences of the IG constraint alongside notation for mutually incompatible plans in a single mind. In chapter four, we record the notation for producing and propagating updates in a chartee's belief structure. In the fifth and final chapter we record advanced moves that one can use to aid the charting process.

We conclude by noting that as an active area of research, optimal charting practice - including even the basic mechanics of charting as discussed here - is bound to change. Indeed, chart mechanics have changed significantly since the inception of CT. See the appendix for a note by Geoff Anders, the inventor of the CT paradigm, on how these mechanics have changed over the course of time.

Acknowledgments

Thanks first to Miya, who both taught us both the material we refined into this document and was the first and closest reader to check it for accuracy. Thanks also to Anja, Emily, Geoff, Jean, and Wolf for enthusiastic pre-release review comments, particularly to Geoff for adding his appendix. Thanks to all the researchers at Leverage and Paradigm who developed the practices this document describes and answered our questions about them, and thanks to both organizations as a whole for supporting our work.

Contents

Chapter 1

Encountering a Chart

This is a document about charts, and it would be a terrible shame if the reader diligently studied chart components before encountering a real chart in its full form. Take a look!

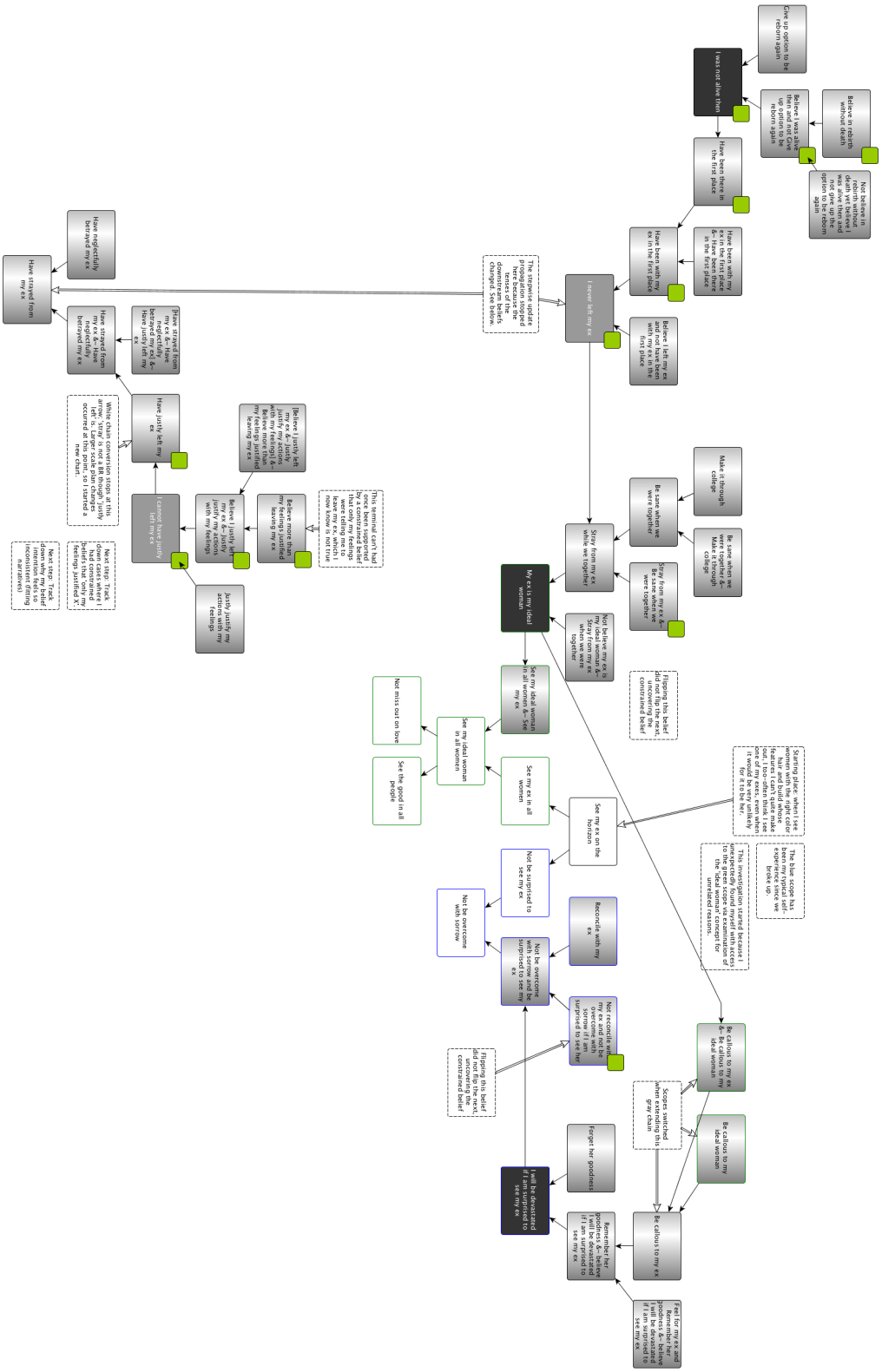

Figure 1.1: A large chart with many interrelated belief structures.

You'll note that there are many pieces in Figure 1.1. We'll point each one out below and note where to look in this document if you want to learn more.

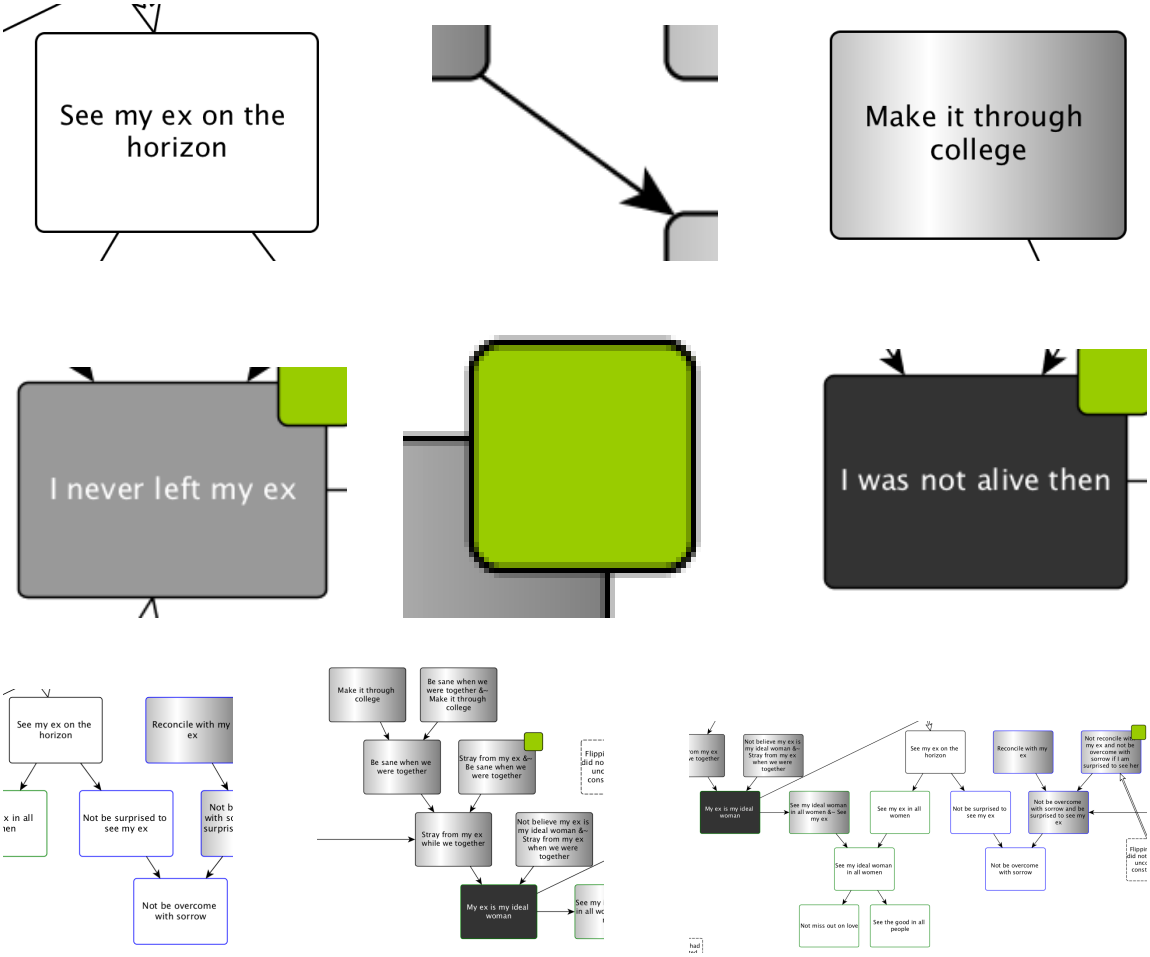

Figure 1.2: The components of a chart.

In the first row from left to right, *white boxes* record actions the chartee plans to do (section 2.1). *Descending arrows* record logical relations between box content, so that an update to a belief in one box will lead to an updated belief in the next one (section 2.2). *Grey boxes* record actions which the chartee does not plan to do (section 2.1).

In the middle row from left to right, grey *attainment response boxes* record beliefs about the world that the chartee must believe to believe that their intrinsic goals are achievable, i.e., are fixed by the CT IG constraint (section 4.2). Green *Update boxes* record which beliefs have been updated in the course of a charting session (chapter 3). Black *injected belief boxes* record beliefs that the chartee believes they must believe due to the CT IG constraint (section 4.3).

In the bottom row from left to right, *white chains* track sequences of actions that the chartee plans to take (section 2.3). *Grey chains* track sequences of actions that the chartee believes they can't take, such that flipping a belief of one box will flip the entire chain below it (section 2.4). Colored borders record *scopes* and *modes*: parts of the chartee's plans which are prima facie mutually incompatible (section 4.5).

Chapter 2

White and Grey Chains

In this chapter we introduce the most common belief structures one finds in charting: white and grey chains. These structures are sequences of actions that the chartee plans and does not plan to take, respectively. We begin by introducing white and grey boxes, the building blocks of white and grey chains.

2.1 White and Grey Boxes: Planned and Unplanned Action

White Boxes

Intuitively, a white box contains an action that the chartee believes they will take. More precisely, the box text is a *belief report representation of an action the chartee can consider performing. The box being white indicates that the chartee believes they will take the action.* We'll unpack this a bit more below.

A box containing the action description "X" will be denoted $[X]$. Pictured in Figure 2.1, the content of a white box is typically obtained by one of two methods:

1. The chartee belief reports that they will take action X. "X" then becomes the content of a white box.

2. After the chartee belief reports that they will take action X, the charter asks the chartee, "What's good about X?". The chartee responds by belief reporting that an action Y will occur as a result of doing X. "Y" then becomes the content of a white box.

White boxes are often read with a "good to" preceding the box content. For example, the top left box could be read "*good to brush my teeth*". This is not always the case - sometimes the belief report is solely of the form, "X will happen". In cases like this a charter would still ask for the purpose of X, but in a different format such as"Why will X happen?", or "For what purpose will X happen?". Finally, actions recorded in white boxes can be actions in a very broad sense, not requiring particular agency on the part of the chartee. For example, it is perfectly acceptable for "Have X be true" to be the (action) content of a white box.

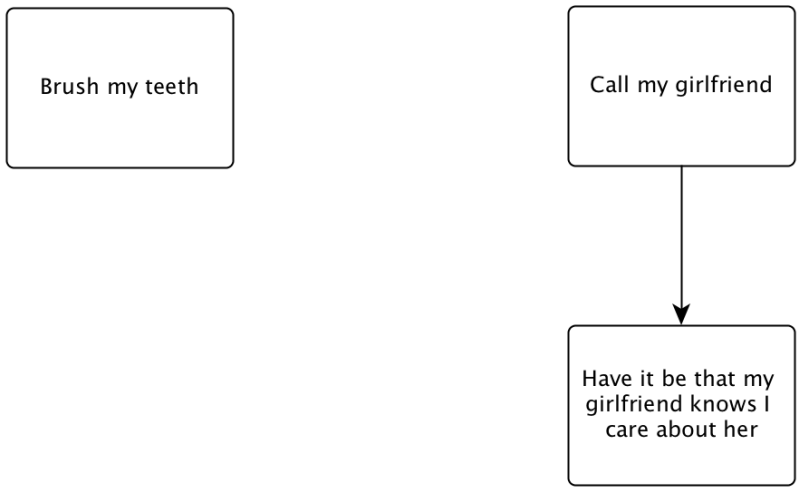

Figure 2.1: At left, we see a white box produced by method 1. At right, the top white box is produced by method 1, and the bottom produced by method 2.

Important features of white boxes

We now unpack our definition of the content of a white box, making a few important notes along the way.

The content of a white box is a belief report representation. Since we are unable to peer into minds directly, a white box contains a representation of the mental content rather than the content itself. The chartee delivers this representation via belief reporting on the mental content associated with the action. The report only corresponds to the chartee's actions to the extent the chartee belief reports correctly, which in practice can be a real obstacle. The content of a white box is an action. White boxes contain actions and nothing else. In Figure 2.1, the box [*my girlfriend knows I care about her*] would be incorrectly formatted since it is not an action, even if it would be a good result of the action in the box above it. The correctly formatted box would specify the action that the chartee plans to take: [have it be that my girlfriend knows I care about her].

The chartee believes they will take the action. This is different from a chartee endorsing doing the action: a chartee can endorse doing something they don't believe they will do, and can believe they'll do something they don't endorse.

Additionally, while they must believe they'll take the action, this belief need not be rational. The chartee may even never end up taking the action in question. This may occur because the chartee has bizarre world models, is subject to psychological constraint, or is ignoring large bodies of evidence. These psych issues and more are some of the reasons people benefit from charting in the first place.

The chartee can consider taking the action. This is straightforward, but the chartee must consider the world in which they take the action to evaluate the goodness of acting. What differentiates

white boxes from grey ones is that the chartee evaluates this considered world as good rather than bad, and therefore plans to take the action. This is also why white boxes containing "X" are typically read as "good to X".

Box content is phrased in the idiolect of the chartee. This idiolect can be substantially different from standard terminology, with concept-content containing unexpected conceptual relations. The same words can even mean different things for the chartee at different parts of the chart. It is important to understand the models that the chartee is actually acting from to generate the behavior in the chart - these are often very different from those assumed by an inexperienced charter.

Grey Boxes

Grey boxes are just like white boxes except that the chartee believes that it would be bad or impossible to do the actions contained within. They therefore do *not* plan on executing said actions. *Precisely, the content of a grey box is a belief report representation of an action the chartee can consider performing. The box being grey indicates that the chartee believes they will not take the action.*

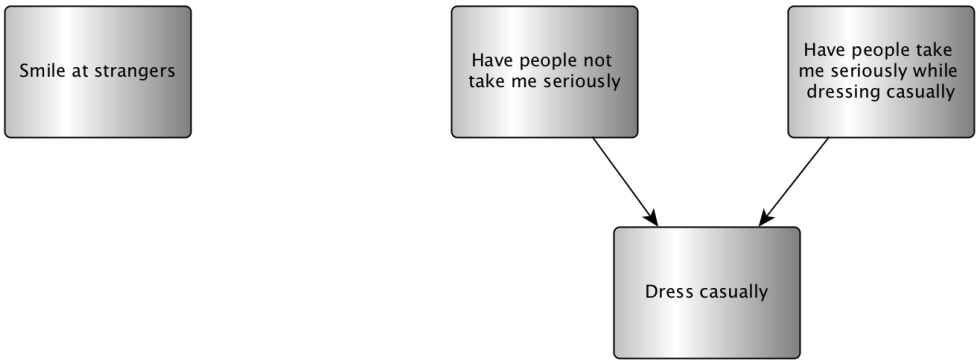

Figure 2.2: The pictured chartee does not plan on dressing casually or smiling at strangers. The rightmost boxes would be read "can't dress casually" because "can't have people not take me seriously" and "can't have people take me seriously while dressing casually". We will explain the arrows and configuration of boxes in the next section.

Important features of grey boxes

Grey boxes are often read with a "can't" in front of them, just as white boxes are often read with a "good to" in front of them. But this "can't" should not be interpreted strictly in every case. According to the definition, "[can't] do X" should be read that the chartee does not plan on taking the action X, even though they could consider performing the action. The chartee does not plan on taking the action because the considered world is evaluated as bad. For this reason, grey boxes are also often read "[bad to] X".

Otherwise, all important features from the above section on white boxes apply to white boxes as well. Grey boxes must contain belief reports about actions and not just assorted things the chartee judges to be bad, they must be interpreted as idiolect rather than everyday language, and they may convey bizarre or irrational beliefs.

2.2 Downward Arrows: Instrumental Implication

Definition

Chart boxes signify *belief* content about actions, and arrows signify *relations* between these beliefs. In particular, we write a *downward arrow* $[A] \rightarrow [B]$ for grey boxed or white boxed actions A and B in a chart if the following two conditions hold in the chartee's belief system:

1. Either (a) A causes B or (b) A is necessary to do B
2. If the chartee could do A, the chartee could do B by an action requiring A.

In what follows we write "A instrumentally implies B" to denote both condition 1 and 2, so that we write the arrow $[A] \rightarrow [B]$ if A instrumentally implies B. We also may say that A implies B as an instrument. Since condition 1 holds if either A necessarily causes B or is necessary for B to occur, we will refer to it as the *necessity condition* for downward arrows. Since condition 2 tracks whether or not "flipping" a chartee's belief about the choiceworthiness of A is enough to cause the chartee to do B, we will refer to it as the *sufficiency condition* for downward arrows.

We note that the causal link between A and B in the chartee's mind may be unclear or may be hard to pay attention to, but the chartee belief reports that B will happen after A. It is OK to use downward arrows - and therefore grey and white chains - to track plans composed of pure temporal sequence without clear causal relation in this case, provided the chartee can belief report that "B happens next".

Examples

The role that downward arrows play is best understood through example. Note the two charts in Figure 2.3.

In the white chain at left in Figure 2.3, we see that for each linked pair of boxes $[A] \rightarrow [B]$, the chartee believes that A causes B (in context). For example, the chartee believes that eating an apple causes them to become appropriately energized, and that competing well in basketball will necessarily demonstrate their excellence. In white chains, we note that it is typically condition 1.a) from the definition of instrumental implication that applies.

In the grey chain at right in Figure 2.3, we see that for all linked pairs $[A] \rightarrow [B]$ in the grey chain, the chartee believes that A is necessary to do B. For example, the chartee believes that going to the movies necessitates eating popcorn, and that eating popcorn necessitates getting sick. The arrows also indicate sufficiency, so that if the chartee believed it was OK to eat popcorn, they could go to the movies, or that if it was OK to get sick, they could eat popcorn. But since the boxes are grey, they do not plan on doing A and therefore do not plan on doing B - for example, they do not plan on getting sick so they do not plan on eating popcorn. In grey chains, we note that it is more commonly condition 1.b) from the definition of instrumental implication that applies.

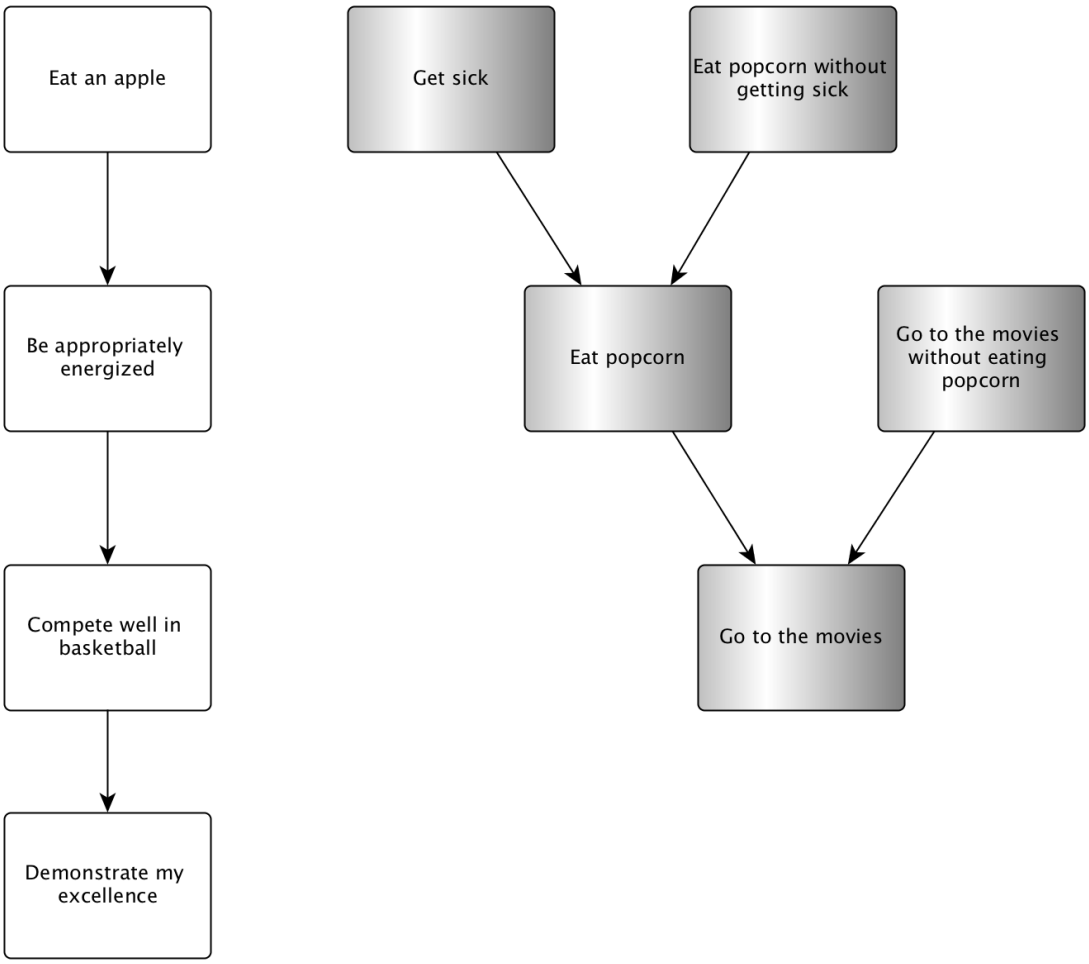

Figure 2.3: The chart at left is a white chain. The chart at left is a grey chain.

We also note that instrumental implications can be *believed* without being *true*. Going to the movies does not force you to eat popcorn. Much of the work of charting is finding and untangling unnecessary instrumental implications and false conceptual conflations in the chartee's mind.

2.3 White Chains

Definition and Example

A white chain is a sequence of white boxes linked by descending arrows. They track a sequence of actions that the chartee plans to take in which each one partially causes the next to occur. We call the top box of a white chain the *head* or *start*, the bottom box the *root*, and intermediate boxes either *nodes* or *boxes* plain. CT posits that one acts at all times to achieve one's basic goals; therefore if CT is correct, one will find a basic goal at the end of any white chain provided one extends the root far enough - no matter how mundane the head.

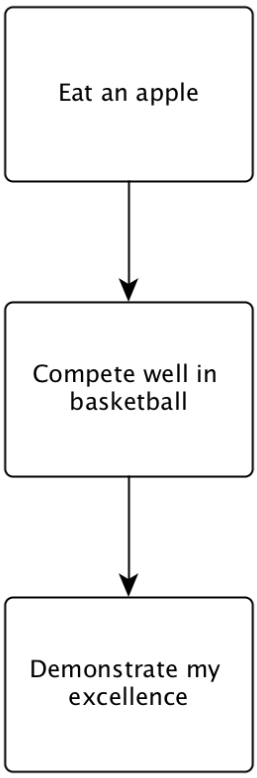

Figure 2.4: A white chain we saw in the previous section. The current root does not appear to be an IG.

Provided you have not bottomed out at an IG, each box in the chain is done instrumentally to obtain the next box in the chain - not for its own sake. Therefore if we have a white chain $[A] \rightarrow [B]$, but the chartee finds an action C which accomplishes B in a more elegant way than A, then the white chain will switch to $[C] \rightarrow [B]$ and the chartee will stop doing A.

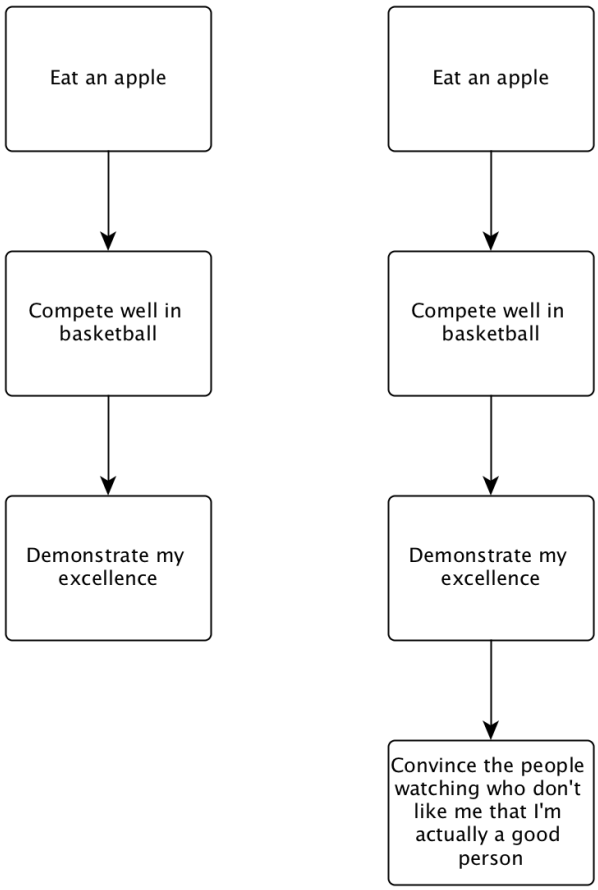

Figure 2.5: The chain on the right is extended from the one at left.

Operations on White Chains

Extension

One extends a white chain $[A] \to \cdots \to [B]$ by finding the action C accomplished by doing B, turning the whitechain into $[A] \to \cdots \to [B] \to [C]$. This is typically done by asking the chartee, *"What's good about B?"*, although it is wise to run checks afterwards to ensure that B is indeed only instrumental to C - for example by asking some variation of *"If you had C, would you stop doing B?"*.

Refinement

One refines a white chain link $[A] \to [B]$ by examining the mechanism by which A accomplishes B. This can be done by straightforwardly asking the chartee, *"How does A cause B?"*. When

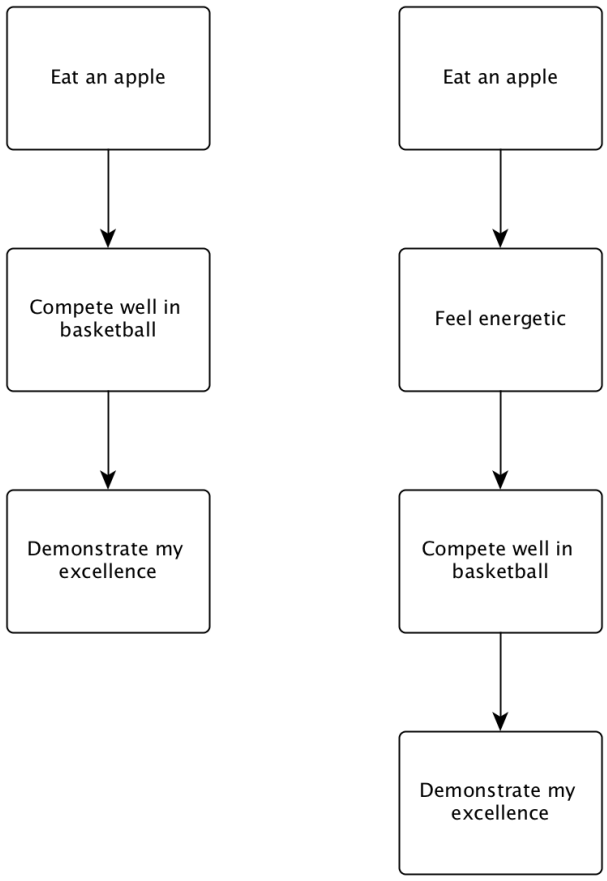

Figure 2.6: The unclear step [Eat an apple] → [Compete well in basketball] in the left chain is refined to the one at right.

the mechanism is not obvious to the charter, there is often interesting psychological content in the believed mechanism, including metaphysics, constraint, and/or perceived powerlessness.

Further Notes on White Chains

Extension as Refinement to IG

If we have a white chain $[A] \to [B] \to \cdots \to [R]$ where $[R]$ is the current root, CT posits that $[R] \to [I]$ for some intrinsic good I, where the arrow might involve a long and complicated mechanism. Therefore extending the chain $[A] \to [B] \to \cdots \to [R]$ to a deeper root $[A] \to [B] \to \cdots \to [R] \to [R']$ can be viewed as refining the chain from $[A] \to [B] \to \cdots \to [R] \to [I]$ to the chain $[A] \to [B] \to \cdots \to [R] \to [R'] \to [I]$.

Transitivity of Instrumental Implication

If the white chain is well-constructed, instrumental implication is transitive: if $[A] \rightarrow [B] \rightarrow [C]$ then one should have $[A] \rightarrow [C]$. This may be an update for the chartee to recognize, rather than being a believed implication from the start, and sometimes such updates are enough to resolve issues on their own.

However, instrumental implication is a conditional implication – a statement of the form "if A, then B" relative to the mental context of the belief report. Therefore transitivity from $[A] \rightarrow [B] \rightarrow [C]$ to $[A] \rightarrow [C]$ depends on whether or not the arrows $\rightarrow [B]$ and $\rightarrow [C]$ have the same conditional antecedents. One can guarantee this by ensuring that the chartee continues to refer to similar contexts and alternatives throughout the white chain. In cases where the transitivity does not hold, the intransivity may itself constitute an important update for the chartee. More about this can be found in the section on chain zoom (section 5.1).

2.4 Grey Chains

Definition

A *grey chain* is a binary tree of grey boxes linked by instrumental implication arrows. They track chains of actions that the chartee does not plan to do.

The Basic Unit of a Grey Chain

The basic unit of a grey chain is a three-box tree - see left and middle in Figure 2.7. At the bottom is a grey box $[A]$, and above it two boxes:

1. A box $[B]$ containing an action that the chartee does not plan to do, but must do in order to do A. We call this the *main line box* of the basic unit.

2. A box [A without B] indicating that the chartee does not plan to do A without doing B. We call this the *side line box* of the basic unit.

See the examples in figure 2.7. Note that both boxes must obey the sufficiency condition, such that *if* the chartee could either do "B" or "A without B", the chartee could then do "A". It is important to not get this wrong, and you can ensure this happens via various checks. For example you can ask the chartee "If you could do B, could you do A?"

Examples

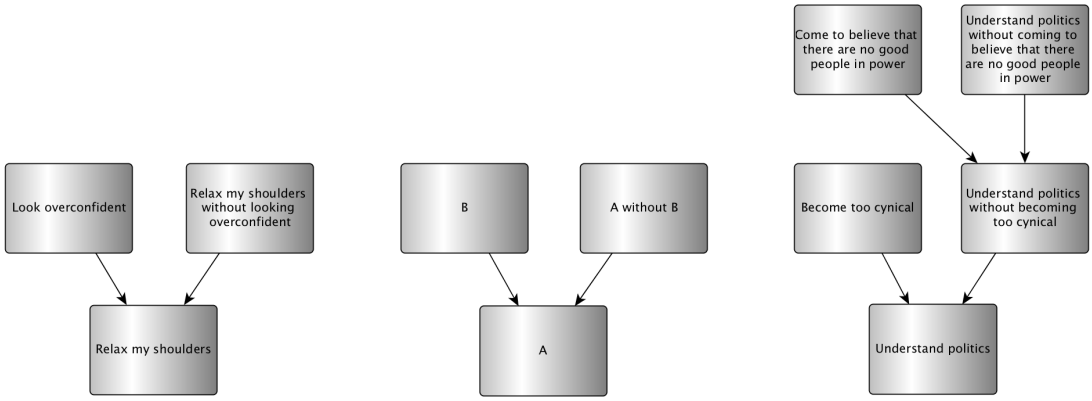

Figure 2.7: On the left, an example of the basic unit structure of a grey chain. In the middle, the general form of the basic unit. On the right, a longer grey chain created by iterating the basic structural unit.

Longer grey chains are simply copies of the basic unit chained together. Terminology for longer chains follows that of the basic unit – the main line of a chain is the sequence of the main line boxes starting from the bottom, and the side line of a chain is the sequence of side line boxes starting from the bottom. In the rightmost example in Figure 2.7, the main line would be

$$[\text{Become too cynical}] \to [\text{Understand politics}]$$

and the side line would be

$$[\text{Understand politics} \ldots \text{power}] \to [\text{Understand politics} \ldots \text{cynical}] \to [\text{Understand politics}].$$

We call the bottom box of a grey chain the *root* and the top boxes with no arrows pointing to them *leaves*.

Operations on Grey Chains

Main line extension

For any grey chain leaf $[A]$, one can elicit the believed reason why the chartee cannot take the action A and then record this in action format as a box $[B]$ lying above $[A]$. This is done by asking questions such as *"what's bad about A?"* or *"why can't you A?"*. We refer to the appending of a box $[B]$ to the main line as *main line extension*.

As an example, take the middle grey chain in Figure 2.7. The chartee reports that it would be bad to be overconfident because then his crush would think he looks stupid. He also reports that he'd be fine looking overconfident if his crush didn't think he was stupid. The charter then extends the mainline of the grey chain – see Figure 2.8.

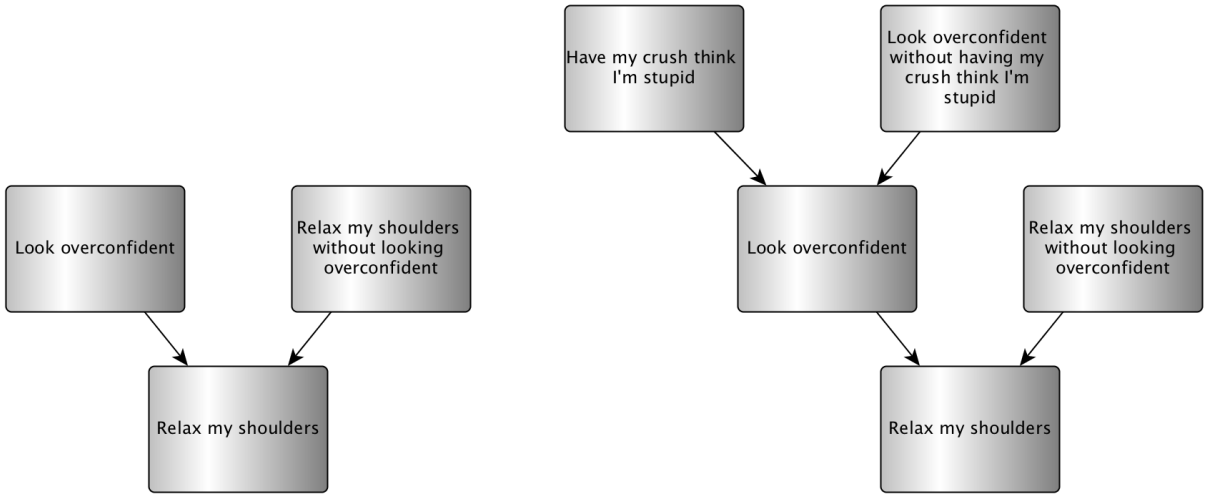

Figure 2.8: On the left, a grey chain. On the right, the same grey chain after mainline extension.

A practical note for charters: when creating a mainline $[B] \rightarrow [A]$, it is important to ensure the chartee is reporting instrumental implication correctly. There is a sense in which B is the "key" or is "crucial" to doing A, and it's worth checking for this sense in the chartee. One failure mode is where the chartee gives "reasonable" reasons B for why they can't A, rather than the crucial reason why they're avoiding A. Another typical failure mode is one in which the chartee says they can't A "because it feels bad". Empirically, charters have found that there is always more substance behind avoidance than things just feeling bad.

Side line extension

$[B] \rightarrow [A]$ shows up in the grey chain because B is viewed as necessary to do A. However, often the mental mechanism behind the believed necessity of B to do A is irrational. In order to track these mechanisms, one places the box [A without B] above $[A]$ any time one does main line extension $[B] \rightarrow [A]$. We call the box addition [A without B] $\rightarrow [A]$ side line extension.

Just as one can extend the grey chain up the main line, one can also extend the chain up the side line. Typically one asks *"Why can't you A without B?"* in order to understand why B is believed to be necessary to do A. For example, take the chartee who didn't want to go to the movies because she wanted to avoid eating popcorn. When asked why she can't go to the movies without eating popcorn, she belief reports that she can't resist the temptation to eat popcorn at the movies. By doing so, we reveal the mechanism of necessity – internal temptation – behind the mainline sequence [Eat popcorn] $\rightarrow$ [Go to the movies].

Sideline extension with respect to grey chains can be seen as an analogue of refinement for white chains since both aim to understand the mechanism by which a believed evaluation of an action B

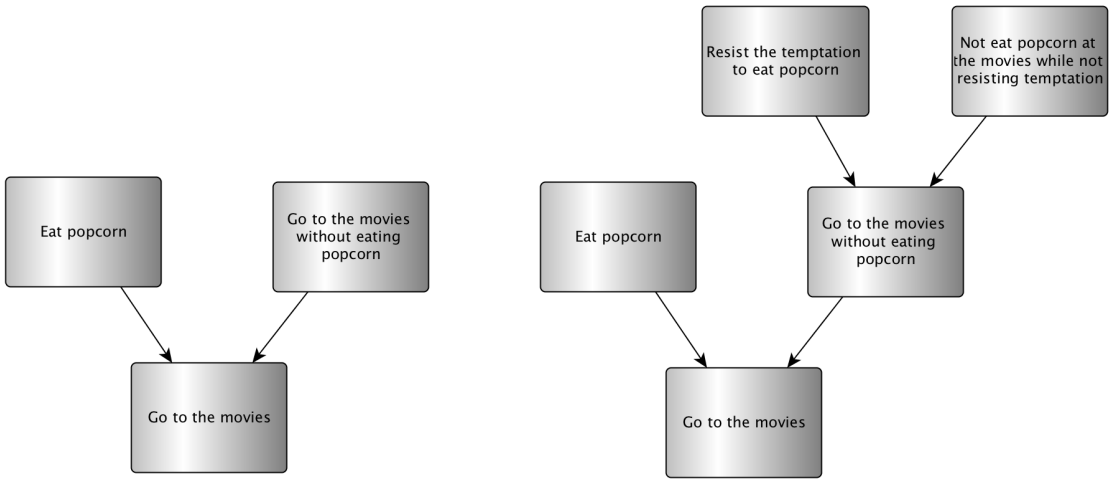

Figure 2.9: The grey chain undergoes side line extension from left to right.

generates a believed evaluation of an action A.

2.5 White-to-Grey Chains

Consider a white chain link $[A] \to [B]$. Since A is solely done instrumentally to accomplish B, this white chain has an "implied grey box" which states that the chartee does not plan on doing B without doing A. By custom these boxes are not always written down in the course of charting a white chain, but they are always implicitly present. See Figure 2.10.

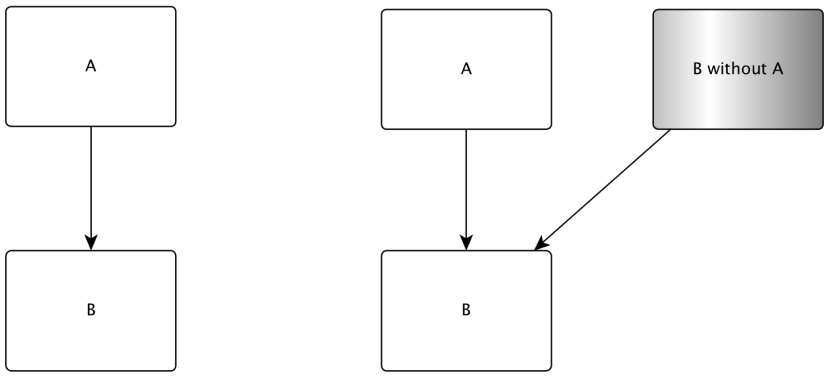

Figure 2.10: On the left, a standard white chain. On the right, a white chain with the "implied grey box".

Despite the implied grey box along a white chain link $[A] \to [B]$, it does not have to be that the chartee sees absolutely no alternative than A to obtaining B. He may see alternatives but then evaluate them as worse options than A - for example because doing them would conflict with achieving other goals.

While the implied grey boxes are not written down in standard practice while charting a white chain, they are often uncovered if one refines the $[A] \to [B]$ chain and attempts to understand the mechanism by which A accomplishes B. This mechanism may be flawed or constrained, in which case it is fruitful to investigate the grey chain with the implied box as the root as in Figure 2.11.

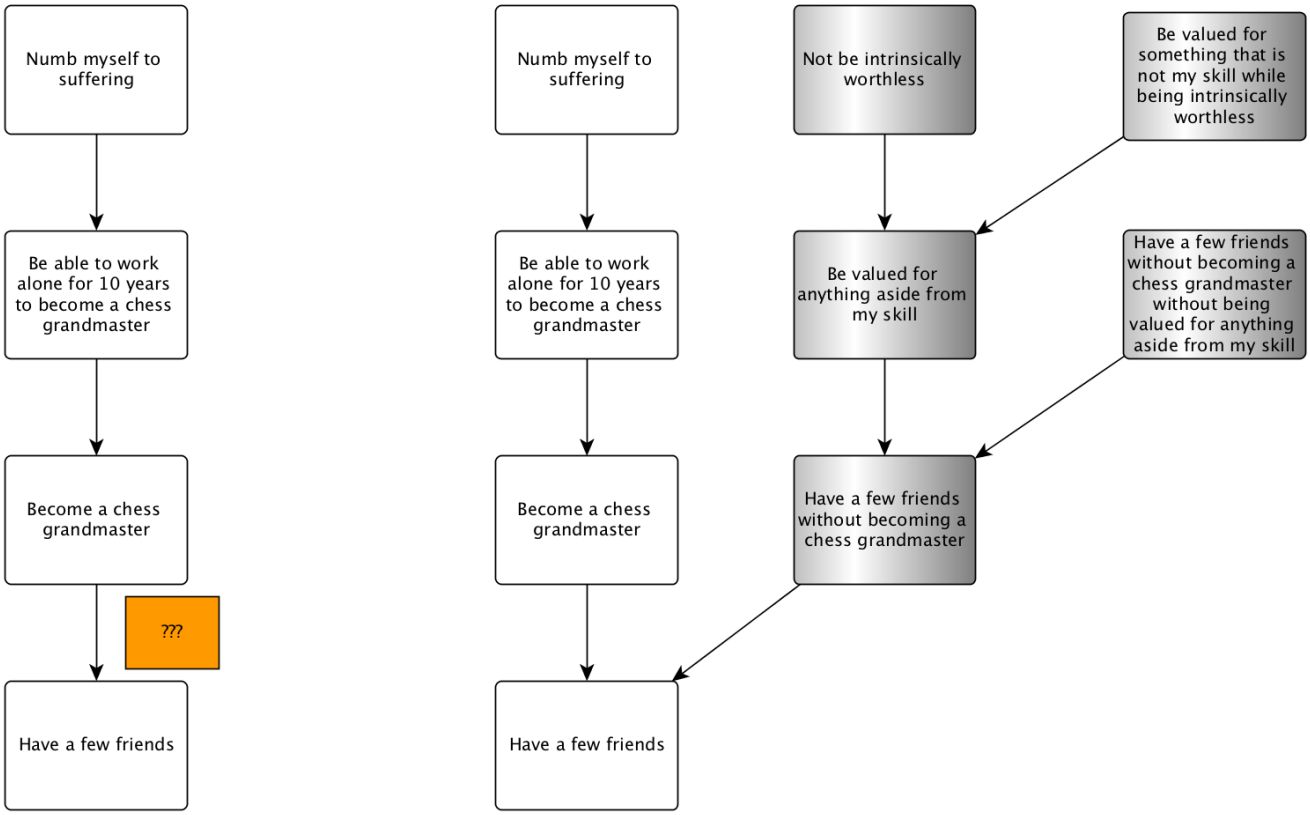

Figure 2.11: At left, we find a chartee becoming a chess grandmaster in order to have a few friends. Sensing that this path was formed subject to constraint, the charter investigates the implied grey box along the last link in the chain and finds significant psychological content.

Chapter 3

Recording an Update

One typically pursues charting to produce targeted belief change in the chartee. In this chapter we describe both the operational and notational process of producing a belief update in a chartee, and "propagating" this update throughout the chartee's belief system to change the belief at the root of the belief structure charted. We will see that propagation is made possible by the logic of implication arrows.

3.1 Updating a Chain Leaf

Recording Leaf Updates in the Chart

If one successfully produces an update in a charting context, the chartee will belief report "[ok to] X" or "[can do] X" where they previously belief reported "[bad to] X" or "[can't] X". Most commonly, this occurs when the chartee belief reports a new grey leaf box and then realizes that the corresponding belief isn't true – the action is actually possible and acceptable. To record this in the chart, place a small green box in the corner of the grey box pertaining to the updated belief. It is also common and advisable to note what produced the update on the chart; see Figure 3.1.

Note: constraints on belief change

According to CT, the mind updates locally and elegantly where it pays attention to explain sensations given the IG constraint. Therefore, if CT is correct, one will update the grey-boxed belief "[can't] X" to the green-tagged belief "[ok to do] X" if:

1. The belief is unconstrained, and
2. The belief "[can X]" is more elegantly explanatory than "[can't] X" with respect to the sensations,
3. "[can] X" is a locally available belief.
4. The attention is placed on evaluating the possibility of X.

There is much to say about each requirement on this list, and there are similarly many strategies one can take to ensure each requirement is met. However, we will confine ourselves to noting

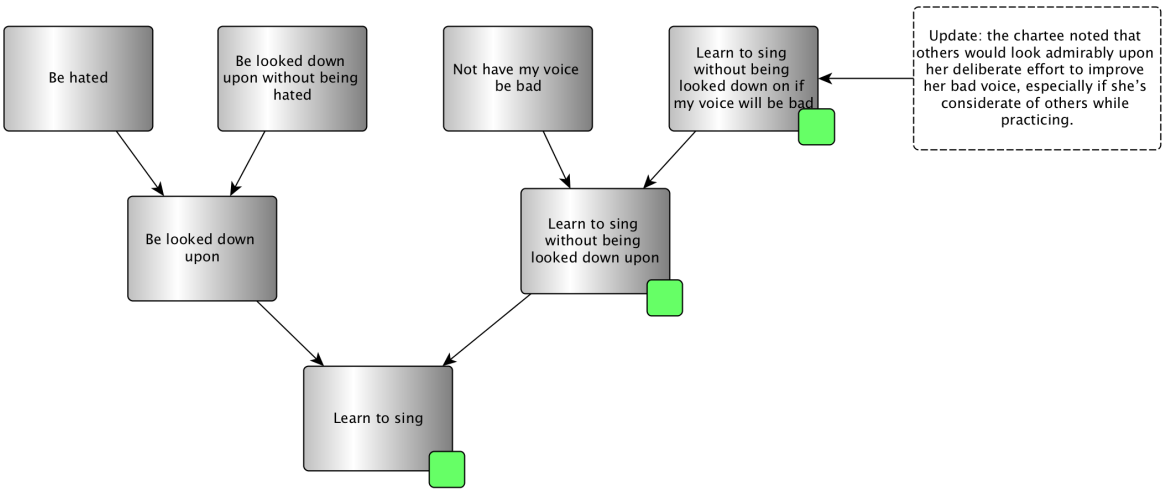

Figure 3.1: Example. A chartee is unable to learn to sing and reports the above grey chain. The charter produces an update at the top-right box.

these constraints on belief change without going into detail about how to cause belief change. The purpose of this document is not to teach the charting moves, but rather to explain the core logic and mechanics of charting.

3.2 Stepwise Propagation

If done correctly, the logic of the chart allows one to "propagate" updates: updating a single grey box A will make it possible to update the grey box below it $[B]$. To see this, recall the sufficiency condition for arrows, which states that the arrow $[A] \to [B]$ implies that if the chartee could do A, the chartee could do B through some action requiring A. By merit of the update, the chartee now believes they can do A. Therefore, placing their attention on the possibility of B and the new evidence that A is feasible, the chartee should become able to belief report that they can do B.

This is not instantaneous, however. To propagate the update, one encourages the chartee's attention sit on the causal relation between A and B alongside the fact that they can now do A. This can be done for example by asking them, "Now that you can do A, can you do B?" and running them through thought experiments in which they have to do B while having them keep in mind that they can do A. In the case from the previous section, the charter might ask the chartee if they can now sing without being looked down upon now that they know they'll be admired for trying to improve their bad voice.

Just as before, one records the result of a stepwise propagation update with a small green box. See Figure 3.2.

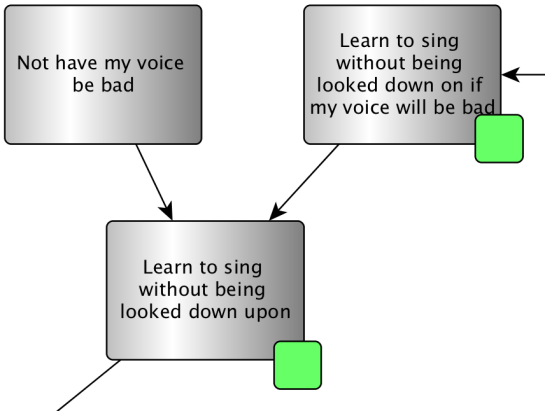

Figure 3.2: After successfully propagating the initial update to the box immediately below, the charter records the update with a small green box.

Handling Propagation Errors

In real cases this propagation can and will go awry - after asking "Now that you can do A, can you do B?", the chartee may respond "No" with a new objection. In that case, the right thing to do is give yourself credit for partial progress, and then copy and restart the grey chain from the node at which propagation failed.

For example, the initial update could have slid in, but when asked "Can you now learn to sing without being looked down upon?", the chartee might respond "No". One then copies and restarts the grey chain. When asked "Why can't you learn to sing without being looked down upon?", the chartee notes that "My dad will be disappointed that I chose singing over piano". One would then proceed to chart that issue. See Figure 3.3.

We reiterate that if the beliefs are recorded correctly in the chart - in particular the sufficiency condition for arrows is obeyed exactly - one will not encounter this issue. In practice, however, this does not occur 100% of the time. In these cases copy & restart is the best way to make progress.

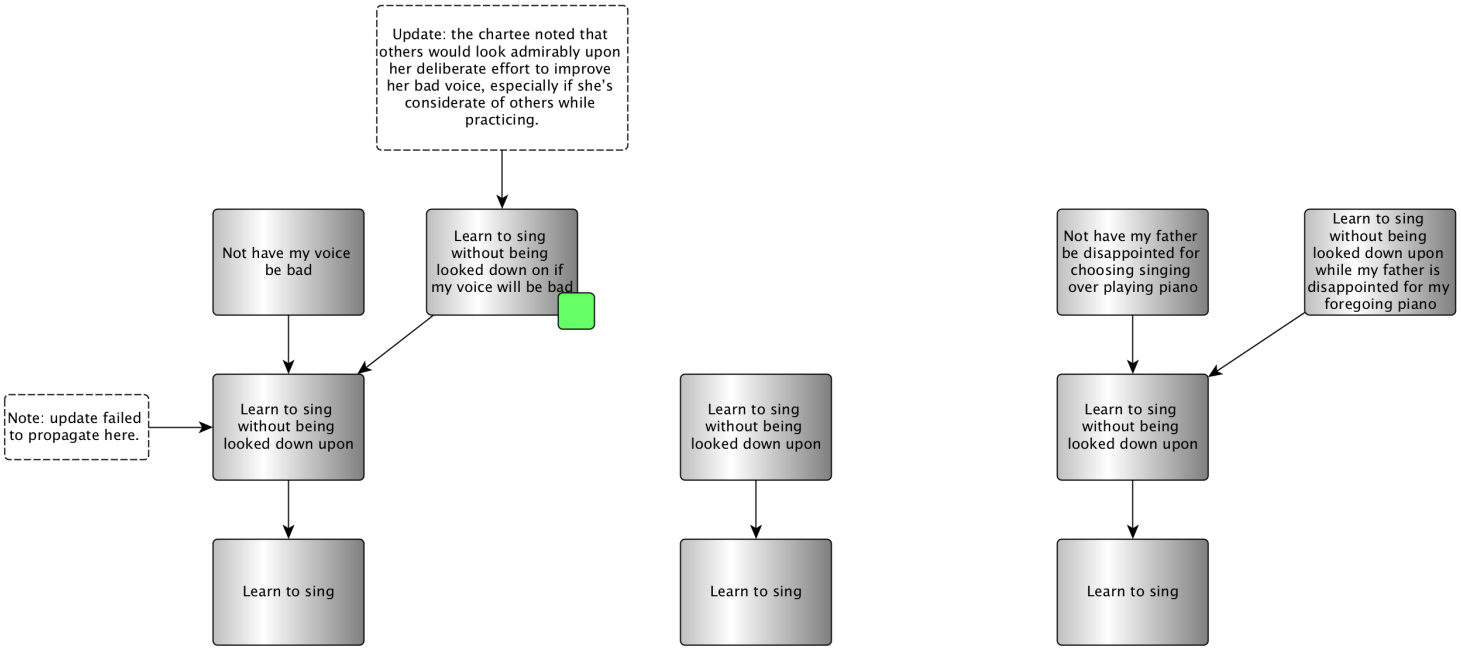

Figure 3.3: The example discussed above. On the left, failed propagation. In the middle, the charter copies and restarts. On the right, charting the new issue.

3.3 Updating a Chain Root

If one propagates an update all the way down to the root, the chartee will report that they *can* do the root. However, they may not go so far as to say they *will* do the root unless there is a real motivation for the root action. A charter should check this once the root has updated. If there is, then the root action is both believed possible and desired, so it will be planned, and the plan for doing the root will consist in the sequence of actions tagged by green boxes. In other words, the root becomes a white box and the chain of updated beliefs becomes a white chain.

In this successful case, some charters choose to color the new chain white while others keep it green-tagged-grey. Some charters find it useful to check in one session later, to see if the chartee is doing the green-labelled actions along the grey chain. To be especially sure, they may white chain the topmost updated grey box. We note that the boxes in the new white chain may contain words that are different from the words in the old grey chain, since the mental content to which they refer has updated under recent attention. It is in this sense that thinking about recently updated mental content feels "fresh", as if there is "something new" about it.

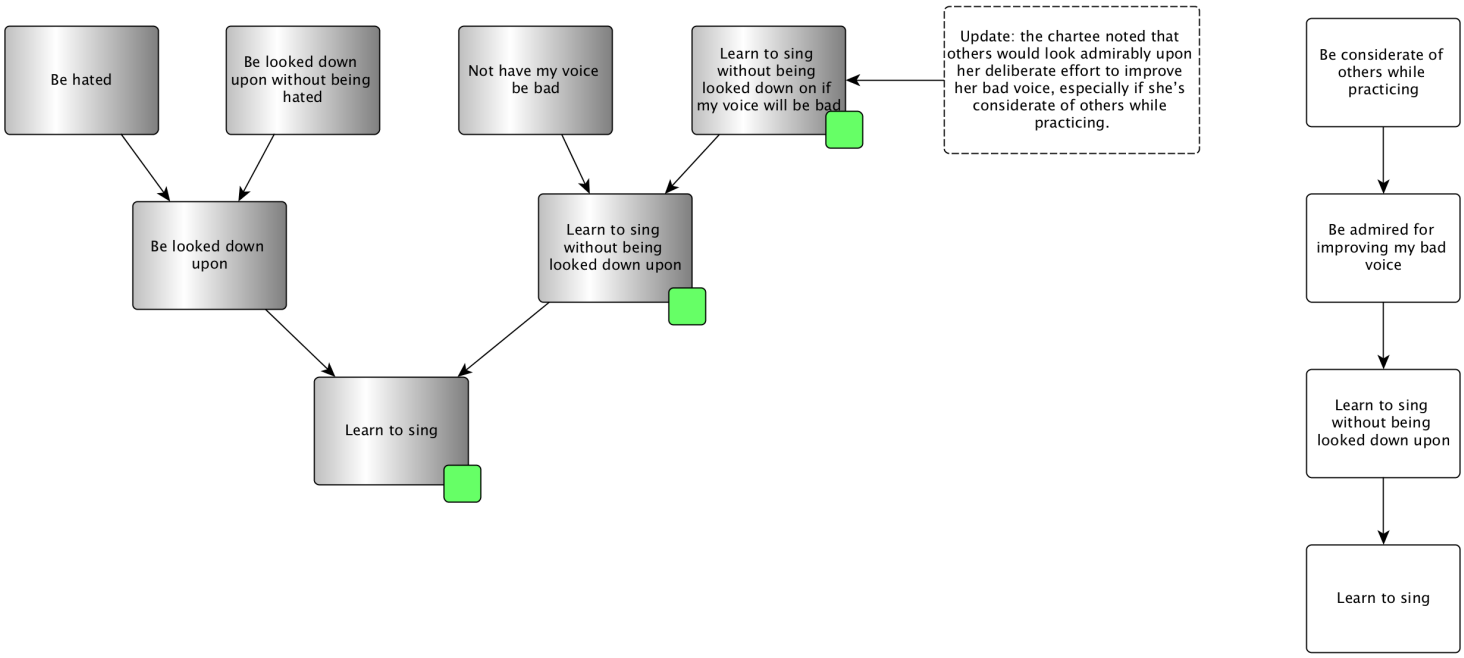

Figure 3.4: A fully updated grey chain alongside its corresponding white chain. Note that the conceptual content of the white chain is slightly different from that of the grey chain due to the content of the updates.

Chapter 4

Constrained Belief and Context

CT posits that minds must believe their basic goals will be stably achieved. This chapter is devoted to the belief structures which arise as a consequence of this assumption. Attainment responses are beliefs mandated by the IG constraint, while injected beliefs are beliefs mandated by the IG constraint alongside a chartee's self-concept. Scopes and modes are prima facie incompatible plans and plans that relate them, which are often formed via constraint to achieve seemingly mutually imcompatible IGs.

4.1 Horizontal Arrows: Analytic Implication

Analytic Implication: Plans Necessitated by Beliefs

In the previous chapter we discussed both inability and planned action that arose instrumentally to achieve goals. There is another way that plans and inability arise: by being conceptually necessitated by beliefs in the mind. For example, if you do not believe that objective truth exists you will be unable to form a plan to find objective truth. When a plan (or lack of plan) P of this sort is *necessarily* believed because of some belief B, we say that B *analytically implies* P.

When these beliefs in question seem irrational, we can try to update them by paying attention to contrary evidence. However, if these doubtful beliefs are not responsive to evidence, we add an unresponsive belief box to the chart and write a *horizontal* arrow to indicate that it analytically implies the plan content of a box. Unresponsive beliefs come in three primary forms: attainment response, injected belief, and interfering plans. These will be discussed in that order in the further sections of this chapter.

Example

We illustrate analytic implication with horizontal arrows in Figure 4.1. In these examples, we see that for each pair of horizontally linked boxes $[A] \rightarrow [B]$, the chartee believes that A necessitates B (in context). For example, on the left the chartee believes that he cannot realize Samantha has flaws because he must believe she is flawless. On the right, the charter notes that if the chartee believes he cannot not create distance, then he must be planning to create distance. In the left case, the charter will then grey chain the impossibility of giving up the belief that Samantha is flawless to investigate how to have the chartee realize Samantha has flaws. In the right case, the charter

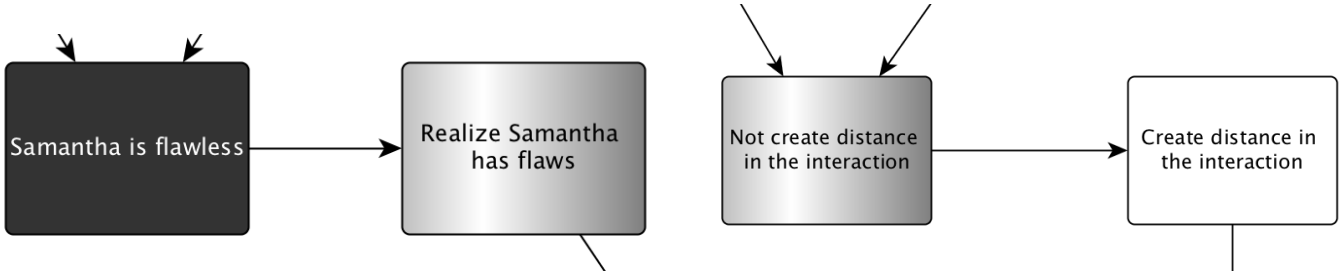

Figure 4.1: Left: the plan to carry out an action X seems impossible because of an injected belief Y. Right: planning to do X analytically necessitates planning not to not do X and vice versa; the conflicting white box is necessary given the grey box.

will white chain the plan to create distance to investigate how to make not creating distance appear possible.

4.2 Attainment Responses: Unchangeable Belief

Definition

In order to believe their basic goals will be achieved, people may be constrained to believe non-IG propositions about the world. Such beliefs may be unresponsive to evidence, and one may belief report that it would be bad if the belief in question were not true. We call these beliefs attainment responses or constrained beliefs.

Attainment Responses in the Chart

An attainment response is denoted in the chart by a belief, not an action, in a dark grey box - see Figure 4.2. As usual, the belief is obtained via a belief report. Instead of reading the constrained belief box containing the belief "X" as "[can't] X", one reads the box as "[can't have it be other than] X".

One arrives at the attainment response via a white or grey box which is analytically necessitated by a constrained belief. For instance, when asking why an action A is impossible, the chartee may respond that a feature of the world, rather than an action, makes it impossible even though there is no such feature of the world. One then forms a grey chain with the constrained belief X as a root, beginning with the query "What would be bad about X not being the case?".

Example

Stockholm syndrome is an example of an attainment response. Say that John forcefully kidnapped Kim and kept her in isolation from all people aside from himself. The police find and jail John and release Kim from captivity. When being interviewed by a psychologist, Kim finds herself unable to recall bad details about John and finds it impossible to imagine John possessing flaws.

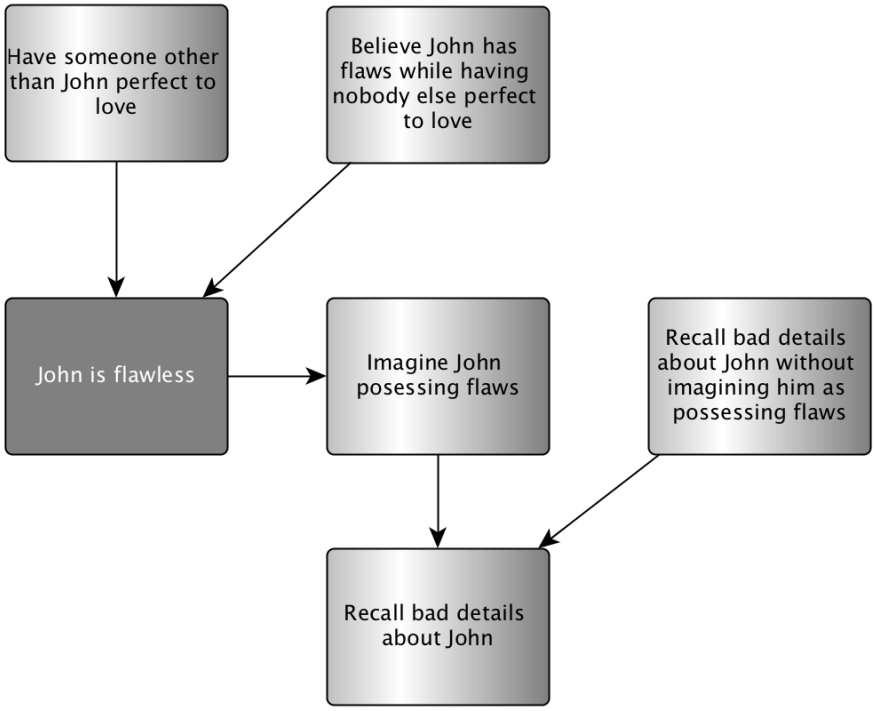

Figure 4.2: Kim's attainment response case.

She then belief reports that John in fact is flawless, and that it would be critically bad if John were not flawless. When asked why, she responds that if John were not flawless, there would be no way to find someone else perfect to love since she only had contact with him.

There are two important principles illustrated in this example. First, one typically arrives at constrained beliefs by charting grey chains and noting analytic necessity rather than noting the constrained belief itself. Finding a constrained belief directly from a white chain is rare, although one often does find them via implied grey boxes alongside white chains. Second, one notes a drastic separation between the issue originally charted (recalling details about someone) and the issue lying above the grey chain (shape of path to connection IG). This is typical of constrained beliefs, and is one reason why they're signified by a special box and horizontal shift in the chart rather than just another grey box labeled "[can't] believe X".

4.3 Belief Injection: Planned Belief

Definition

People may believe that *they need to believe* certain propositions in order to achieve their basic goals. These beliefs too are often unresponsive to evidence, in which case a chartee will belief report that it would be bad if they didn't believe the proposition in question. We call these beliefs

constrained belief injection.

Injected Beliefs in the Chart

An injected belief is denoted in the chart by a *belief*, not an action, in a black box. As usual, the injected belief is obtained via a belief report. Instead of reading the injected belief box containing "X" as "[can't] X", one reads the box as "[can't believe other than] X".

One arrives at the belief injection response via a white or grey box which is analytically necessitated by a constrained injected belief. One then forms a grey chain with the injected belief X as a root, beginning with the query "What would be bad about not believing X?". One typically finds an injected belief by charting grey chains and noting analytic necessity, and typically it will be the root of a main line chain very different from the grey chain that led to the injected belief in the first place.

Example

Rick is dating Samantha, who isn't the best girlfriend. Despite thousands of dollars of debt accumulating and causing immediate problems in his life, Rick can't seem to pay attention to the fact that Samantha is mismanaging the couple's finances - let alone bring himself to manage them himself.

After consulting a charter, Rick belief reports that it would be bad to pay attention to the situation since it would be bad for Samantha to have any flaws. This analytically necessitates that he currently believes Samantha does not have any flaws. After checking the attainment response with a belief report, he reports to the charter "It would be bad if I didn't believe Samantha was flawless, because then I'd leave her, and then I'd be alone forever". See the chart in Figure 4.3.

Contrasting attainment response and belief injection

The difference between attainment responses and injected beliefs is subtle. In both cases, you have a constrained belief B. The difference arises from the type of constraint: in the case of an attainment response, B must be true about *the world*, otherwise the chartee would not achieve their basic goals. In the case of an injected belief, B must be *believed in the chartee's mind* since otherwise the chartee would not achieve their basic goals. When charting attainment responses one often finds issues with world models, while when charting injected beliefs one often finds issues with self-concepts.

We see this in the examples from the attainment response and belief injection sections of the chapter. Both resulted in the chartee's constrained belief that their romantic partner was flawless, but the fear underlying the constraint was different. In Kim's case, the fear was of a *world* in which connection was impossible. In Rick's case, his fear was of *what would happen if he believed* that Samantha was flawed, namely his acting to cause connection to be impossible.

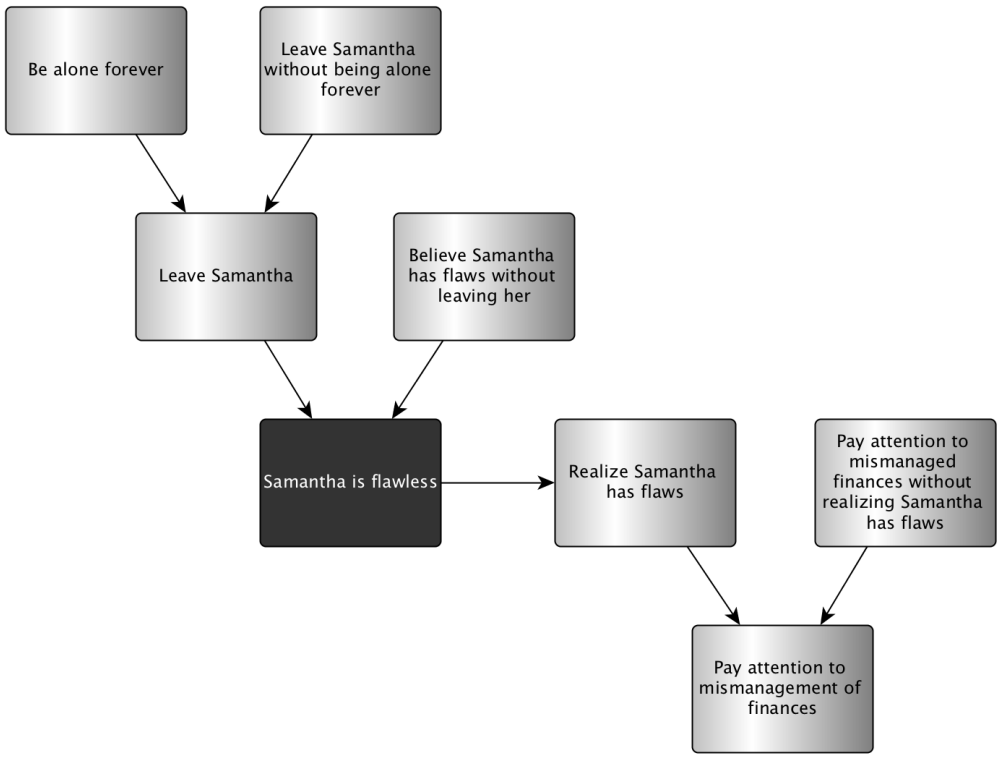

Figure 4.3: A chart displaying Rick's belief injection.

4.4 Grey-to-White Chains: Plans Implying Other Plans

Introduction

It is sometimes useful to reframe a non-action as an action, or an inability as an ability. For instance, in *volitional powerlessness*, the mind will willfully avoid performing actions or unjustifiedly believe in its inability to do actions in order to achieve some positive purpose. In the closely related case of *inverted self-concept*, the mind misconceives a plan to perform an action as a powerlessness not to perform an action.

Grey-to-White Chains in the Chart

This reframing is done on a chart by turning a grey box into a white box, in which "[can't] not X" becomes "[good to] X". Since these boxes are related by analytic necessity, we format the chart so that the boxes are connected by a horizontal rather than vertical arrow. One can then investigate what's good about being *unable* to do an action.

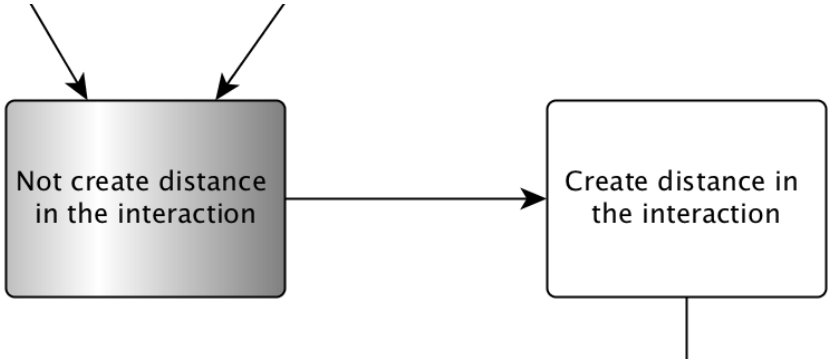

Figure 4.4: Not planning to do X analytically necessitates planning to not do X and vice versa.

Examples

Case 1: Bad endorsement. In Figure 4.5, a particularly ambitious chartee comes to the charter with the "problem" that he isn't working 18 hours a day. When asked what would be bad about working 18 hours a day, the chartee reports that he would not see family and friends. If he did both, he would have to sleep 4 hours a night. *"Hmm,"* the charter thinks, *"maybe this isn't the update we should be going after"*.

The charter then constructs a white chain tracking *what's good* about not working 18 hours a day and finds that the chartee is not working 18 hours a day in order to live a harmonious life. In constructing a white chain from a grey chain, the charter allows the chartee to reconsider a bad endorsement.

Case 2: Volitional powerlessness. The mind will often willfully avoid performing actions or unjustifiedly believe in its inability to do actions in order to achieve some positive purpose. In cases like these, it is useful to chart the good thing achieved by the powerlessness.

In Figure 4.6, a fully grown adult with adequate cooking models belief reports that it would be bad for them to make dinner for themselves. When asked about what's good about not cooking for themselves, they report that doing so will force their romantic partner to cook for them, which in turn will demonstrate that they are loved.

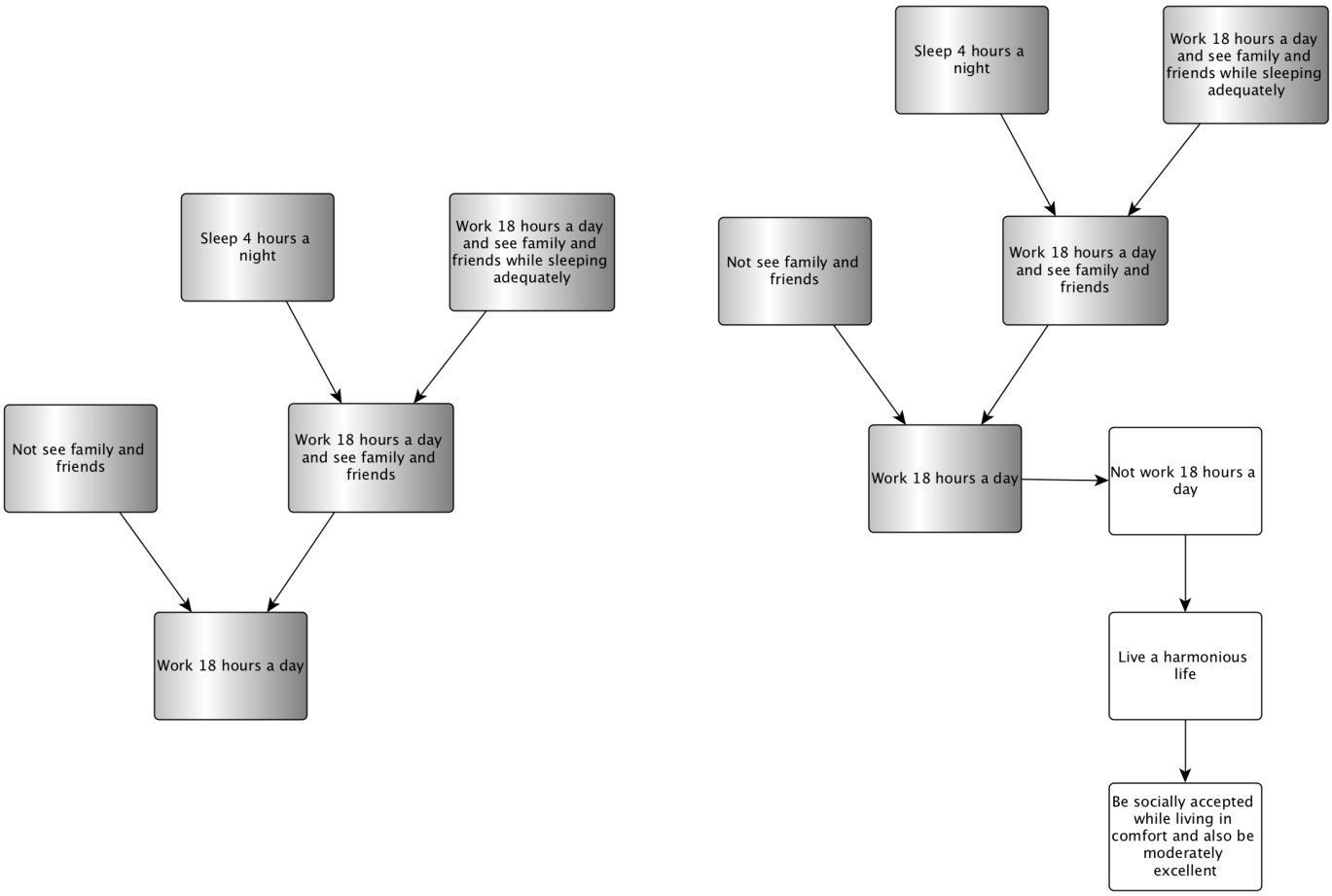

Figure 4.5: Endorsements are produced by particular models and/or IPs and are flawed insofar as those models or IPs are. One should not always pursue the initial endorsed update presented by the chartee.

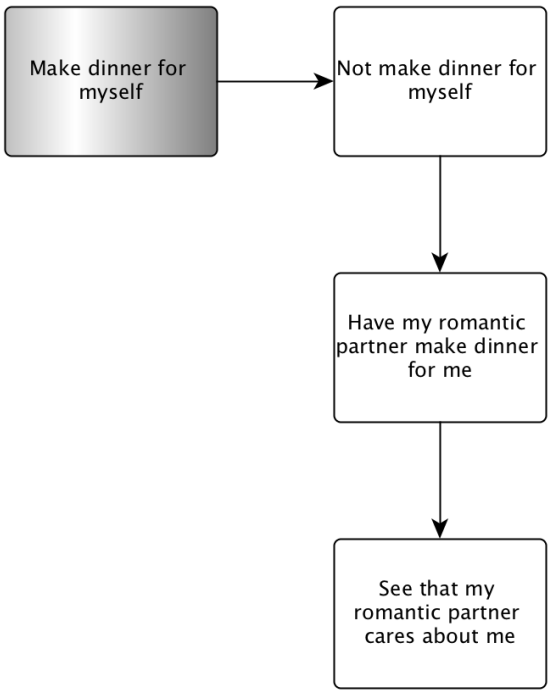

Figure 4.6: An example of volitional powerlessness.

4.5 Scopes and Modes: Plans to Change Plans

Introduction and Definition

While charting, we find that people report seemingly inconsistent statements. Sometimes this is a problem with belief reporting or with attention-management IPs, but not always. People in fact do switch back-and-forth between apparently mutually-inconsistent plans in their minds, switching based on context that the chartee may not be aware is guiding their local plan choice. In such cases it is useful to track this context switching using box border colors to indicate which implicit context the chartee makes each belief report from.

When the contexts are fully absorbing, so that the chartee only introspects within one at a time exclusively and the chartee does not feel the context-dependence from within, we call these *modes*. When the contexts are simultaneously open for introspection or the chartee is aware of how the multiple contexts relate, we call these *scopes*.

Since context-dependent plans are plans just like others, they are represented on a chart by a collection of white chains, grey chains, or attainment responses as we have already discussed. To track mode or scope information these otherwise identically interpreted boxes are also given *colors*

to indicate the mode or scope context. That context shapes the meaning of the text's idiolect and the nature of the counterfactuals used to evaluate the value of each action in the text.

Example

We take this example from Connection Theory - Theory and Practice. Consider Bill, who often finds himself switching between two sets of behaviors. Often he will work himself into a frenzy while fantasizing about the great things he will accomplish, and he often informs his father of the achievements he makes during these episodes. These episodes may last for many days or weeks. Whenever he sees his slacker friends, however, he suddenly feels lazy, just wants to chill, and will typically avoid his father. After asking a charter for help, he finds the following apparently inconsistent belief structures in the two different episodes: see Figure 4.7.

The plans in Figure 4.7 are mutually incompatible, given the first requires having a strong work ethic and the second requires being a slacker. Since Bill does not act on and report these simultaneously, this is classed as a mode difference rather than a scope difference; if Bill learns to understand the mode structure in the course of charting to the point of understanding both as part of his larger plans then the mode difference will become a scope difference.

The color records that the charter recognizes that Bill can only believe both "it is good to be a slacker" and "it is good not to be a slacker" insofar as context can make these noncontradictory. As it stands, there are at least two major context differences. First, the evaluations may be referencing different possible worlds: one in which slackers can rely on society to be kind to them vs. another in which society is callous so that they must rely on hard work and family connections to thrive. Second, the relevant concept behind the words "be a slacker" may have changed: one concept by which "slacker" means a person who spends time having fun and another by which "slacker" means a person who squanders opportunities for achieving great things.

Context-Specific IG Paths

According to the belief rule in CT, minds must believe at all times that their basic goals will be achieved. Therefore if CT is correct, each mode and each scope will have a path to every one of the person's basic goals. Some of these paths will be particularly inelegant based on persistent concepts the person holds. For example, if Bill has a model of romantic connection via mutual powerlessness, his romance path in the above blue mode may be particularly inelegant since the blue mode requires his becoming powerful. It may be so inelegant that desiring romantic connection switches him out of his blue mode and into a different one.

In this case, risk of not stably achieving an IG in one mode context causes the person with multiple modes to change their assumed context to one in which that IG will be stably achieved. The plan is thus implicitly to choose one's context to stably achieve one's IGs. In and around scopes, a chartee frequently has better developed and known plans for how to manage context switching to ensure the IGs are stably achieved. In the example above, for instance, a charter might lead Bill to develop a plan to be a positive-sense slacker (a person who spends time having fun) when he's off work without becoming a negative-sense slacker (a person who squanders opportunities for achieving great things) while he's working to ensure he stably achieves both acceptance by his father and social acceptance.

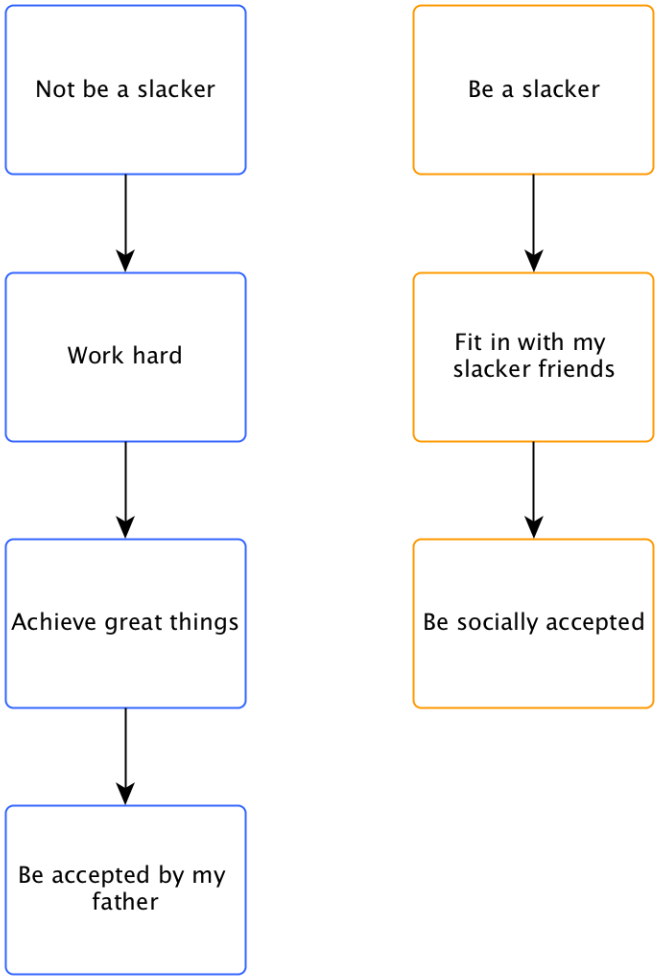

Figure 4.7: Example: a zoomed-out mode chart.

Unpacking Context

Because of the mutual incompatibility of modal and scoped plans, one cannot do inference or logic between modes or scopes unless one understands the metacontextual interface between the two plans. *Trigger beliefs* - beliefs which cause one mode or scope to switch to another - are one such interface. *Context-specific IPs* – IPs which maintain the elegance, consistency, etc. of believing that one is and acting as if one is in a particular context – are another.

Trigger Beliefs

Two things characterize a trigger belief:

1. It is crucial to both modes in question. That is, whether or not the belief is held determines whether an associated mode is very elegant or very inelegant.

2. The trigger belief is tenuously held, so that it is able to flip and may do so frequently.

In Bill's case, a trigger belief could be his self-perception of power. Bill's holding this belief or not holding this belief will determine which mode is more elegant, and therefore which mode he will find himself believing. The trigger belief is tenuously held, given that neither a strong work ethic nor slackerdom are seriously baked into Bill's self-concept.

One can chart modes without determining the trigger beliefs and determining a trigger belief can be hard. Given knowledge of Bob's behavior in the example chart alone, the trigger belief could be:

1. importance of paternal social acceptance,

2. low capacity to work hard,

3. whether or not his girlfriend approves of him being a slacker,

or any number of other things. There also may not just be a sole trigger belief, but several or many. Trigger beliefs between scopes may explicitly reference the context shift, for instance if there were a trigger belief "now is the time to work for great things" activating Bill's blue-bordered white chain in the initial example.

Context-specific IPs

Each mode and scope can carry a set of context-specific IPs which serve particular purposes. Examples of IPs and the purposes they serve for are:

1. Attentional redirection away from evidence indicating path inelegance within a given context

2. Once mode switching occurs, manufacturing of imaginations or quick firing of theoretical IPs to "tile out" the rest of the context-dependent models

3. Manufacturing imaginations confirming elegance of the beliefs contingent on the current context

4. Managing the availability of imaginations of counterfactual possible worlds to match contextual assumptions

We certainly are not exhaustive here. Context-management IPs for modes are limited by the strict separation of mode-specific belief content, whereas context-management IPs for scopes are less limited in what they can reference from other scopes.

Chapter 5

Advanced Moves

This chapter describes some advanced moves that are frequently useful in charting contexts. Chain zoom uses transitivity of instrumental implication - when it applies - to distill belief structures. Branch splitting and branch joining allow you to reconfigure grey chains to more easily track disjoint concepts in a single box, or similar concepts along a chain. Concept cleanup allows a charter to produce updates across a belief structure by replacing one concept with another one better-fit to describing reality.

5.1 Chain Zoom: Transitivity of Implication

White and grey chains track believed plans or believed obstacles to plans, which often contain many steps. Chain zooming is the act of condensing the belief structure tracked on a chart into its key believed components. This is possible insofar as instrumental implication is a transitive relation, i.e., if $[A] \to [B]$ and $[B] \to [C]$, then $[A] \to [C]$.

In most well-constructed chains, instrumental implication is transitive, i.e., if $[A] \to [B] \to [C]$ then one should have $[A] \to [C]$. However, instrumental implication is a conditional implication: a statement of the form, "if A, then B" relative to the mental context of the belief report. Therefore transitivity from $A] \to [B] \to [C]$ to $[A] \to [C]$ depends on whether or not the arrows $\to [B]$ and $\to [C]$ have the same conditional antecedents. One can guarantee this by ensuring that the chartee continues to refer to similar contexts and alternatives throughout the white chain. In cases where this transitivity does not hold, chain zoom is **not** valid, and trying to check chain zoom checks this transitivity (e.g., asking *"if you could C, could you also A?"*).

Example: Chain Zoom when Transitivity Holds

Consider the example in Figure 5.1, taken from a personal chart. I found myself in the course of a psychological investigation unable to condense the material I had written into a few simple questions, despite endorsing that this was possible. After going up a relatively long grey chain I found that I couldn't write the questions down because then I would realize how long it would take to solve the psych issues described by those questions. If this occurred, I would realize that I wouldn't improve fast enough, and then would get left behind by the ecosystem.

The main benefit of valid chain zooming is to understand the structure of the psychological issue at a larger scale without being bogged down in the details. This can give a chartee a better

understanding of what's going on in their mind, alongside allowing the charter do diagnosis and find the main bottlenecks. In my case, the simple structure of the issue couldn't be directly read off from the original chart, and was only understood by chain zooming.

Examples of Intransitivity

The coarse-grained belief structures arising from chain zoom are not always plausible, since they can represent links between widely separated particular beliefs. Sometimes the evidence of this implausibility is enough to cause an update in its own right, as when a chartee realizes that their plan is far too convoluted or magical, laughs, and switches to a plan to find a more plausible plan. More commonly intransitivity is not so simple to address.

Intransitivity can occur either because of errors in the chart or because instrumental implication is a conditional implication. The line between error and modality can be tricky to specify exactly, so it is worth considering several representative cases.

Case 1: change in idiolect-content matching. When a chartee belief reports, the mental content represented by their words depends on their representation intention. For instance, some chartees will find themselves unable to belief report with certainty about other people if they are concerned that people should not be represented statically. If a chartee's intentions change during charting, then later boxes may correspond to a changed idiolect compared to earlier boxes. If this happens at a single step, one may then find intransitivity of instrumental implication across the box or boxes where the change occurs. This is because the contents necessary for the boxes below were not identical to the contents sufficient for the boxes above: the boxes where the change occurred have two different referents for the two different layers of arrows.

Case 2: change in saliency of alternatives. When a chartee considers whether an action is possible or good, they do so against a background of various other possibilities. Which particular possibilities seem relevant may change based on interaction with the charter or sustaining attention on a destabilizing topic. When these possibilities change, then the plausible instruments and requirements for action may also change, altering the background against which descending arrows take their meaning and possibly causing inconsistencies between arrows of different layers of the chart.

Case 3: change in assumptions about the world. Changed belief about which possibilities for action are plausible and salient is a special case of changed belief about the world in which action takes place. When assumptions about the world change in the course of charting, but in a way that does not affect idiolectical representation of earlier box content, the chart boxes may still seem correct while the arrows cease to follow transitivity. For instance, early boxes may assume the chartee is in the charting location while later boxes may assume a hypothetical in which the chartee has travelled back to their home country, so that a newfound ability to do something thought impossible in a later box will not be sufficient to enable downstream actions unless the new ability applies in the home country specifically.

In cases such as these it is often useful to identify, explicate, and update the idiolect, saliency, and context models directly in order to help clarify the chartee's thinking and planning in the area. Updates of this kind will commonly correspond to concept cleanup, and so chain zoom is a natural way to look for opportunities for that.

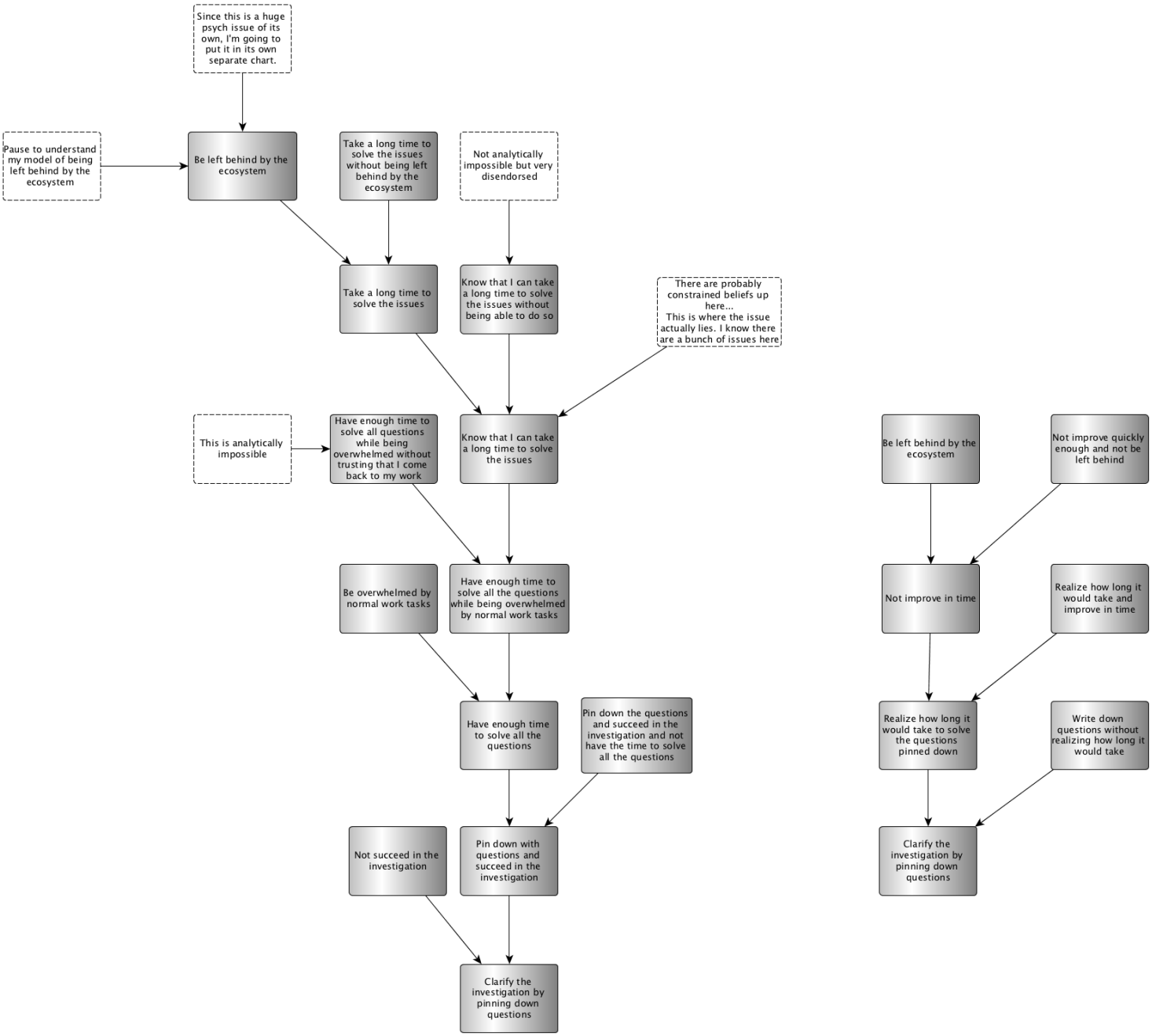

Figure 5.1: At left, the original chart. At right, the chain-zoomed chart.

5.2 Branch Splitting and Joining

Branch Splitting

Say a chartee believes that *either* B or C would be sufficient to get A, but can't A - and therefore can't B or C. For example a chartee can't be cool because he can't buy nice clothes, nor can he become persuasive - see Figure 5.2.

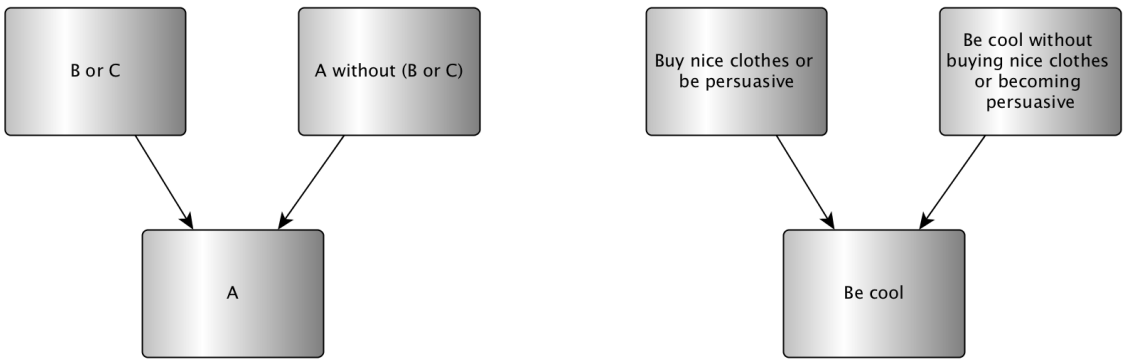

Figure 5.2: The grey chain corresponding to the above scenario.

It can be cumbersome to grey chain [A without (B or C)] if B and C are conceptually distinct in an important way. Fortunately, we may transform the grey chain in a way which notationally separates B and C. We call this *branch splitting*, since it amounts to splitting one branch into two. See Figure 5.3.

The split-branch chart allows you to chart the two conceptually distinct issues B and C along separate grey chains. This is particularly useful when a chartee has reported a conflation of two concepts that the charter wants to split apart.

Branch Joining

In the example above, we took a single concept "B or C" and expanded it into two separate grey chains on the chart since "B" and "C" were conceptually distinct. The reverse move is also useful - when two grey chains "B" and "C" are conceptually very similar, one can condense the two into one by grey chaining "B or C". A toy example of this is provided in Figure 5.4. We call this branch joining.

In Figure 5.4, a chartee reports that they can't eat a snack. They can't eat a snack because they can't eat an apple. They don't have an alternative to that since they can't eat a banana. They don't have an alternative to that since they can't eat an orange, and so on. The charter condenses the entire grey chain into one box, "[can't] eat an apple, banana, or orange". After confirming with a belief report, it turns out that the chartee doesn't want to eat a fruit of any kind. They replace the several similar branches with a single joined branch then proceed normally.

It is worth noting that the difference above is not just notational. Branch joining can lead the chartee to belief report more concisely and more comprehensively, where before the counterfactuals

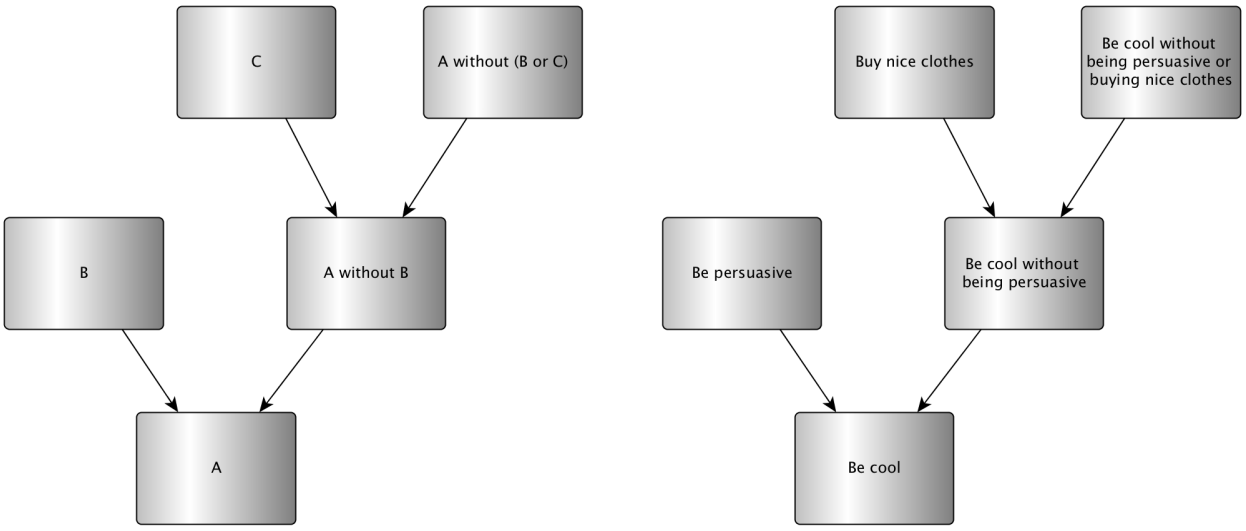

Figure 5.3: The branch split grey chain.

to their grey boxes were leaking out into side line chains. This is very useful for the charter and chartee.

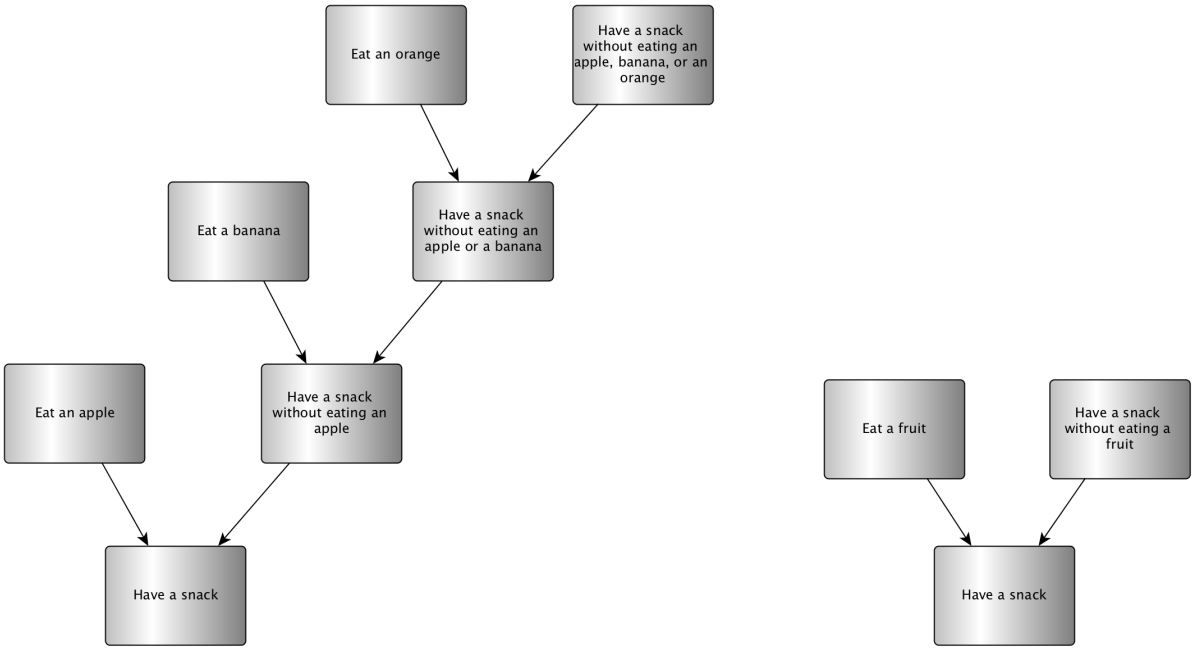

Figure 5.4: An example of branch joining a grey chain.

5.3 Concept Cleanup: Reassigning Variables

Definition

This document has discussed the standard means of producing updates in a chart, namely, change a single grey box to white via elegant constrained updating and propagate the update down a grey chain. In this section we discuss a different form. When charting an issue, there may be a single concept involved in many components of the grey or white chain rather than just one. *Concept cleanup* is the act of updating the ambient concept and thus updating all the boxes of the chain that reference that concept thereby.

Example

For example, consider Tim, who is working hard in school to become excellent so that others might respect him. When the charter asks Tim to elicit his understanding of respectable excellence, Tim comments that he hasn't really thought much about it before. After a minute or two of pondering, Tim exclaims, *"Wait - now that I think about it, I'll only be respected if I become a genius, and school won't help me with that! It's actually sucking up time that I could use to develop my own theories."*

In Figure 5.5, we see that Tim's natural updating function caused a change in the concept of excellence which was lying in the background of the action chain. This new concept of excellence then caused him to adopt a new plan, updating the entire chain at once. This need not always

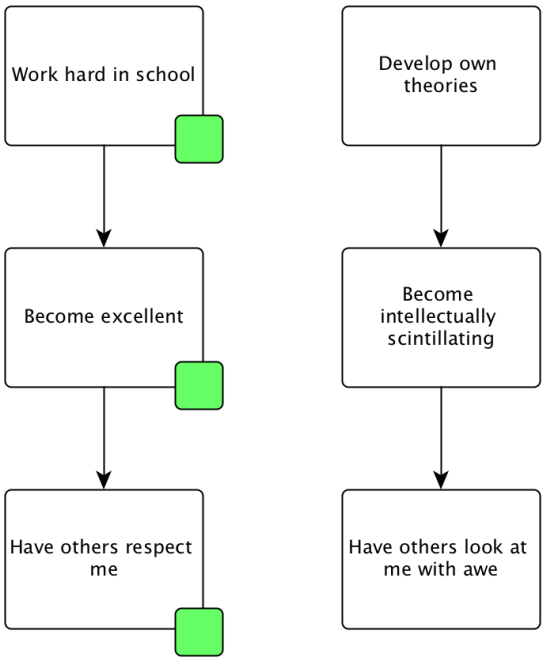

Figure 5.5: At left, Tim's old white chain. At right, Tim's white chain after swapping out his concept of intelligence.

occur when cleaning up concepts. It may be that the main structure of the path remains, but some aspects of the intermediate steps shift. It could also be that the path remains intact but downstream psych issues created by inelegance of the original background concept are knocked out.

The charter should note a few things about concept cleanup in addition to the above. First, in this example the concept swap was produced by Tim's attention and the natural updating function alone. This need not always be so - often the charter will explicitly construct an alternative concept and hand it to the chartee to see if it slots in. Second, the above example shows that the path structure and idiolect involved in the chain may change after doing concept cleanup. Therefore when one notes an update in a concept reused often in a grey or white chain, it is prudent to re-check key components of the chain. Finally, note that refinement of chains and concept cleanup are complementary, since they respectively examine and shift mechanisms and models which are lying in the background of a sequence of belief reports.

Appendix A

Changes in Chart Mechanics Over Time

A.1 A Note from Geoff Anders

CT chart mechanics have changed over time.

One important notable change is the role of white chains. Original CT charts sought to display all of the important sets of jointly sufficient conditions for a given white box. Current CT charts seek to display only cruxes. This difference gives rise to a number of downstream differences between original CT charts and current CT charts:

1. Original CT charts used both And-arrows and Or-arrows. And-arrows were the default and used regular arrowheads. Or-arrows were used occasionally and had a white circle in the place of a regular arrowhead. Current CT charts only use Or-arrows, which have regular arrowheads.

2. "White chains" in original CT charts typically had a tree structure rather than a chain structure. Sometimes the structures were more complex. White chains in current CT charts typically have a chain structure.

3. Original CT charts would often connect white boxes that were highly useful to one another even if they were not explicitly connected in the person's mind. This was to account for the mind's tendency to develop new links, even if those links did not already exist.