

# Excuses in Philosophy, Psychology and Logics.

Bachelor Thesis  
Jiske van der Lof  
3407918

June 10, 2013

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Philosophy</b>	<b>5</b>
2.1	Philosophy of Language . . . . .	5
2.2	Ethics . . . . .	7
2.3	Philosophy of law . . . . .	9
2.4	Summary . . . . .	9
<b>3</b>	<b>Psychology</b>	<b>11</b>
3.1	What is an excuse? . . . . .	11
3.2	How do excuses work? . . . . .	13
3.3	Negative site of excuse . . . . .	14
3.4	Summary . . . . .	15
<b>4</b>	<b>Logics</b>	<b>16</b>
4.1	Concepts . . . . .	16
4.2	Restrictions for conduct . . . . .	17
4.3	Restrictions for excuse . . . . .	21
4.4	Restrictions for the observer . . . . .	21
4.5	Summary . . . . .	21
<b>5</b>	<b>Integration</b>	<b>22</b>
5.1	Perspectives . . . . .	22
5.2	Restrictions . . . . .	23
<b>6</b>	<b>Conclusions and Discussions</b>	<b>29</b>

# Chapter 1

## Introduction

An excuse is a very common concept. Excuses are seen everywhere in any situation because of lots of different reasons. Doing something you should not, not doing some you should or being caught red handed doing something you should not do, are all situations in which excuses can be made. Since the word excuse can be used in a lot of different contexts it is hard to see what an excuse really is and how it works. Some excuses are evaluated as good excuses, some as bad excuses and others do not even count as excuses. In this thesis I probe what restrictions have to need to be met, so that an excuse is evaluated as an excuse. My aim is to formulate a model that is able to decide if something is a proper excuse or not. In order to make this model, we must first get a well-informed definition of what an excuse is. Different disciplines are interested in excuses. These disciplines all use a different approach to define what an excuse is. Therefore it is interesting to make use of an interdisciplinary approach to form an overarching definition of excuses. The threat of probing a certain concept from only one discipline is that tunnel vision may occur. Important aspects, not noticed in a certain discipline, might be excluded from the final definition. By using several disciplines this risk is reduced to a minimum.

Three disciplines that are of great importance for research on excuses are philosophy, psychology and logics. These three disciplines complement each other. Philosophy will enable me to form a theoretical definition of excuses. Psychology will be of added value since it differs very much in approach from philosophy. In psychology most research is done by experiments or in the fields, rather than applying theoretical research. This gives me the ability to check if theoretical ideas found in philosophy match with reality. The last discipline, logic, is of great value to make the model. The approach used in logic is very abstract. Since it does not only make use of natural human language it can sometimes bring new insights, which are hard to explain otherwise. Some other disciplines could have been very useful as well, such as for example law, but due to the limited time and size of this thesis and my expertise in this field of science I have chosen to limit myself to the three disciplines mentioned above. However, I will use philosophy of law in the chapter on philosophy.

In the first three chapters of this thesis I will discuss excuses from each discipline individually, without already referring to other disciplines. The reason for doing this is that it is not possible to decide in advance which concept of which discipline is most important. Therefore all chosen disciplines have to be used without bias of other disciplines as much as possible. After this process is done the disciplines will be evaluated and compared to one another. This will lead to one overarching definition of the concept of excuses in which the different approaches and concepts of the disciplines are integrated.

Since the concept of excuses is very broad and complex I will apply some restriction to narrow the research field. First of all, I assume that making an excuse is not the same as apologizing for something. With an apology a person confirms he did something wrong and

states he feels sorry for that action. The excuse I want to probe sees to it that a person does not have to apologize. Secondly, some people make excuses before they do a certain action. The aim of such an excuse is to lower the expectations of the observer towards the action of the excuser. The aim of an excuse afterwards, which I want to probe, is not the same; the excuse is made because of something that happened. Therefore other restrictions will hold for such an excuse. Thirdly, the aim of this thesis is not to make a model in which it can be tested, whether an excuse will be accepted or not. This is not possible, since that is too subjective. All kinds of things play a roll in this, such as the mood of the observer, familiarity with the excuser etcetera. Therefore the model will only be able to test if something is a proper excuse. In the chapter on philosophy this will be explained further. By trying to create this model, it is also tested if there exists a logic which is able to describe all restrictions on excuses. Lastly, I assume that the communication during an excuse is normal<sup>1</sup> Thus the excuser and observer do speak the same language and interpret each other's sentences in the right way.

In summary, the aim of this thesis is to find a definition of excuses and translate this definition into a model that shows what is needed for an excuse, which is made afterwards, to be regarded as a proper excuse. The definition is an overarching definition of concepts and restrictions on excuses found in philosophy, psychology and logics.

This thesis deals with questions of AI, as it combines several disciplines belonging to this field of science. By probing excuses I am able to show I obtained expertise in several disciplines of AI. Moreover, I developed my interdisciplinary research skills while combining those disciplines. The interest for AI lays especially in the fact that in this thesis three disciplines of AI are used to probe one concept

---

<sup>1</sup>I found this restriction in an article by Searle [18]. This article is not particular on excuses, but excuses do belong to the category of words argued in his article. I did not want to leave this restriction to philosophy, since I think it is an important presupposition which should hold in all disciplines.

## Chapter 2

# Philosophy

The aim of this chapter is to find a well-formed theoretical intuition of the concept of excuse. Excuses are discussed in language philosophy, ethics and philosophy of law. These three kinds of philosophy all have slightly different approaches and therefore it is interesting to discuss them all. Since the article in which the discussion about excuses starts is written by Austin [1], a language philosopher, I will start with this part of philosophy. Followed by a discussion on how the concept of excuses is approached in ethics and in philosophy of law.

### 2.1 Philosophy of Language

Excuses and justifications often get mixed up, but there is, according to Austin [1] a clear and well-defined difference. There are two main ways to defend a certain conduct. One way is to admit that person A did X, but in doing so he did a good thing, or at least in that context. This way of defending is a justification. The conduct is no longer a bad thing; it was the right thing to do in the current situation. On the other hand, it is also possible to defend a certain conduct by admitting that the conduct was not right, but it was not the agent's fault. This is for instance the case if he was forced, if it was an accident, or if he did not know his action would lead to the result it led to. So in the first defense the agent is responsible but did not do anything wrong. In the latter defense the conduct is wrong but the agent is not responsible, or at least not fully responsible, for the conduct<sup>1</sup>. The latter is what Austin calls an excuse. It may seem that this is a bit of an artificial way of making a distinction. It happens a lot that people use excuses and justifications as defined above together. But still, even when used together it is quite easy to see which part is an excuse and which part is a justification.

Austin makes another interesting but not totally successful distinction, i.e. between locutionary and illocutionary acts. However, this has been criticized by Searle [17]. Austin claims to have a good way of dividing up sentences into two groups: locutionary acts and illocutionary acts. A locutionary act is uttering a sentence that only describes facts or a certain opinion. For example: the grass is green. Illocutionary acts actually contain more than just the plain facts or opinions stated in an utterance. Due to how it is uttered or formulated more information is transferred to the observer. For instance a promise<sup>2</sup>: I will not be home late tonight. This is not just a factual statement; the speaker did not just describe the future, which happens when a weatherman pronounces that it will not be raining tonight. The weatherman does not make

---

<sup>1</sup>I use conduct in this thesis to describe the action done by the excuser for which it might be necessary to excuse. Thus the conduct is the behavior which is evaluated by an observer for if it is good or bad.

<sup>2</sup>Not only promises are illocutionary. All sentences which contain more than just facts or an opinion are illocutionary

a promise; he just shows the outcome of his calculations. The force, which is present in the promise and absent in the weathermen's utterance makes a sentence illocutionary. Austin is sure that locutionary acts and illocutionary acts are mutually exclusive. This is however exactly what Searle refutes. He states that all members of the class of locutionary acts are also members of the class of illocutionary acts, due to the fact that no complete sentence is force free. Since with every sentence the speaker at least shows he knows some facts or has an opinion.

At the same time, Searle agrees with Austin is right on some points. Austin argues that you can look at a locutionary act as an attempt to perform an illocutionary act, whereby the illocutionary is regarded as such when the performance is successful. Searle does not think this distinction will bring much good. Therefore, he does not go much further than to agree with Austin on this point. Nevertheless, I think this might be a very interesting point for this thesis. An excuse only works as an excuse when it is accepted as such by the observer to who the excuse is directed. It seems an excuse is a very delicate concept. It fully depends on the receiver whether the excuse is accepted and leads to the excuser to no longer be held responsible. However, there are some objective restrictions an excuse has to fulfill to be a proper excuse. The last point made by Searle [17] in which he uses the idea of Austin helps to find these restrictions. An illocutionary act is divided into two parts, the proposition, and the force or illocutionary type of the act. Thus the proposition that you were frightened may be a common content of different utterances with different illocutionary forces. The same proposition can be used with different types of illocutionary acts and also the other way around: the same type of illocutionary act can be used with different propositions. Symbolically this can be seen as follows:  $F(p)$ . Where  $F$  determines the range of types of illocutionary forces and  $p$  is a variable over the range of possible propositions. Thus,  $F$  denotes the illocutionary part and  $p$  the locutionary part. To make  $F(p)$  successful, at least the following three conditions have to be met. First, when an excuse has been made,  $F$  is successful when it was the senders' intention to make an excuse and, second, the receiver understood it as such. Lastly,  $p$  is successful when the receiver accepts  $p$  as a proper way of making an excuse.  $p$  is just the utterance of a certain fact.  $F$  is what makes this utterance an excuse.

In a subsequent article [19], Searle continues on this subject. From this article some interesting restrictions can be deduced for a proper definition of excuses. Before it is necessary to make an excuse for a certain conduct, somebody has to accuse the person of that conduct; otherwise an excuse makes no sense. This will be further explained in the chapter on psychology. The purpose of giving an excuse is to no longer be responsible for the conduct somebody accused the person of. Therefore the view of the world of the observer needs to be changed. This has to happen because the person expresses the belief that  $p$ . Thus  $p$  causes the view change and therefore an excuse is an action.

Since  $p$  has a very important task I will further specify this task. The article of Stalnaker [21] about assertions provides some help in this respect. He states that an assertion is the expression of a proposition. Thus, the expression of  $p$  is an assertion.  $p$  represents the world in a certain way. Therefore, what applies for assertions definitely has to apply for  $p$ . First, assertions have to have a context. The speaker has certain beliefs and the intention, in this case, to make an excuse. An observer, who of course has his own beliefs, is essential to make the assertion too. Secondly, the content of an assertion only leads to an excuse if it is made in the right context. A certain assertion is only seen as part of an excuse if an excuse is in order. Third, an assertion is representing the world as being a certain way. Therefore, it is able to change the perception of the world of the observer; it affects the context. In the particular case of an excuse, it changes the beliefs of the observer, which hopefully for the speaker, changes his opinion about the liability of the speaker. The assertion helps to harmonize the beliefs of the

observer and speaker, thereby reducing the set of different possible histories for the observer.

In conclusion, an excuse is an illocutionary act, which can symbolically be seen as  $F(p)$ .  $F$  is the part of the utterance, which makes the utterance an excuse.  $p$  changes the beliefs of the observer. Although the intuition of what an excuse entails is much further defined, some parts are still unclear. The following section will look at an excuse from a different perspective; Ethics.

## 2.2 Ethics

The questions to be answered in this paragraph are: When will an excuse be made? What is expected of the agent making the excuse? What can be excused for? When does an excuse work? These questions are found and answered in articles by Austin, Brandt and Baker. Brandt presents a utilitarian theory of excuses. Although this thesis is not specifically interested in a utilitarian definition, a big part of Brandt's theory is very useful in a more general theory of excuses. Baker responds to Austin's *A Plea for Excuses* by discussing what an accident is. Baker states that the claim "I did it accidentally" works as a powerful excuse because it excludes a charge of negligence. Therefore the assumptions she made about accidents will also at least partially be useful for the concept of excuses in general.

Brandt [5] agrees with Austin [1] that an excuse only occurs when something bad happened, due to the person who excuses. According to Brandt, an excuse works if it clears the excuser of fault. Brandt considers moral excuses. A moral excuse is a good excuse if the person, after he excused, is no longer morally responsible for his conduct. A moral excuse is only needed when someone has done something morally wrong. Brandt makes the same differentiation between excuses and justifications, which is made in this thesis in chapter 2. What is wanted from a theory of moral excuses is described by Brandt as follows: *a set of correct universal statements, describing circumstances in which the agent of an action, which is morally questionable in some way, is properly freed from moral criticism or condemnation or disapproval* [5, p. 340]. This would only be a part of a complete theory according to Brandt. Brandt developed an utilitarian theory of excuses. Since the utilitarian theory states that there is always a good side and a bad side, in all theories both sides have to be discussed. Therefore Brandt has also developed an aggravating theory. Since the aim of this thesis is to create a more general theory, I will not discuss that part of his theory. I will keep to those parts that are of interest for a more general theory on excuses. In the traditional utilitarian theory something is an excuse if the claim made in the excuse proves that it is bad or fruitless to blame the agent for the conduct he committed, even though the conduct is wrong. In addition, it is important that a theory of excuses not only focuses on acts or conduct, but also on which state of mind caused the conduct. The state of mind can be helpful to formulate a proper excuse.

The second theory mentioned by Brandt is a rule-utilitarian theory of excuses. It defines a total excuse as follows: *An agent is morally blameworthy (praiseworthy) for an act if, and to the degree that, the moral code the currency of which in that society would maximize utility would condemn (praise) him for it* [5, p. 353]. This means that an excuse is good when the moral code as described above, no longer condemns the agent after making the excuse. From this, Brandt concludes that somebody can only be excused for a certain conduct, if this conduct does not show any defect of character. It is strange he argues this, since he cannot explain what this actually means. Different options are possible. However, I think the best option is the following: A defect of character means that the person's character is not as it should be according to the moral system. If the conduct shows he has a bad character he cannot be excused for this conduct. Thus, only if the conduct is in conflict with the character of the

excuser, an excuse is appropriate. This matches with the earlier stated idea of Brandt that the state of mind that caused the conduct is of great importance. If the conduct fits the character, then the state of mind is probably different from when the conduct is committed by a person whose character and conduct conflicts.

If you translate these ideas of Brandt to a more general theory of excuses, an excuse has to contain the following constraints: (1) The conduct for which an excuse is made has to be wrong. (2) After the excuse is made the excuser is no longer liable to the conduct. (3) The claim made in the excuse is the cause of the fact that the excuser is no longer liable. (4) The state of mind in which the act is done should be considered in the deciding process whether an excuse is good or bad. (5) An excuse is good if it does not show any defect in conduct.

Baker agrees with Brandt and Austin about the fact that an excuse can only occur following bad conduct. In addition, she brings up two more restrictions [2]. First of all, it is important that a person, in this thesis the excuser himself, is involved in the conduct in some way. Moreover, it is not possible to excuse for something, which happened because of some natural disaster. Furthermore, it is also not possible to excuse for something you are victim of. The second restriction follows from the first. The conduct, which is excused for, might have been intentional in other situations. This means that the conduct should be one which can be intentionally caused by a human being, but does not necessarily has to be.

Doing something purely accidental means that the agent meets all the required standards of care. When examining accidental acts Baker concludes the following: accidents are more like happenings than actions. The only way in which the agent plays a role in the action is that he *unknowingly* provides the context in which the accident occurs. The agent does not initiate the conduct, it happens to him. This however, does not mean that the action, which happened to the agent, was done unintentionally. At the moment of the happening the agent did do some action intentionally, however by doing this action a side effect happened which led to the accident, for example, a reflex because somebody scared the agent. For this reason Baker claims that accidents cannot tell us anything about the agent as a person. It does not make clear how a person acts in a normal situation. This leads her to her final statement of why claiming to have done something accidental is a good way of excusing. She claims that this has the power to show that the agent was not, or at least not fully, in charge of his own conduct. This makes him no longer liable for his own conduct.

The aim of making an excuse is to no longer be held responsible for a certain conduct. Therefore, it is interesting to look at an article of Frankfurt [11] in which he discusses “the principle of alternate possibilities”. This principle means that a person is morally responsible for his actions only if he had the ability to do otherwise. Some philosophers see this as *a priori* true. Frankfurt, however, does not agree with this opinion and states that someone may be responsible for what he has done even though he had no choice of doing something else. Having no other option does not constitute the reason for the chosen action. It may seem natural to see moral responsibility and coercion as mutually exclusive, however it is not. Being coerced to do something does not immediately lead to a lack of moral responsibility. As an illustration, person A wants person B to perform a certain action X no matter what. A will do anything to make sure B is going to perform action X. B cannot do otherwise but perform action X for example because A can change B’s thoughts. Suppose B will perform action X without any help of A. Now it will be unreasonable to say B is not morally responsible even though he has no other option than to perform action X. The only thing that was up to B was if he acted on his own or by help of A, however he could not be aware of this. This brings Frankfurt to the conclusion that although a person could not have done otherwise than performing the act, this



fact is a proper condition of his having done it. But this does not necessarily have anything to do with why he did it. B would have done the same if he had had other options. So the principle of alternate possibilities is wrong when it states that being coerced is enough to be excused for moral responsibility. What is really needed is an excuse in which the excuser claims that he would have acted otherwise if he had had alternative options. Therefore, Frankfurt argues that the principle should be replaced by the following: *A person is not morally responsible for what he has done if he did it only because he could not have done otherwise* [11, p. 838]. This principle should be added to the list of constraints of a general theory of excuses.

This principle makes use of counterfactuals. A counterfactual is an implication. It describes the world in a way that it is different from the real world, and states that something would have gone different if the world was that way [13]. This is exactly what happens with the principle of alternate possibilities. If the person was not forced, thus in a world which is different from the real, he would have acted differently.

## 2.3 Philosophy of law

Westen [24] defines excuses in law by making a distinction between *actus reus*, justifications and excuses. This distinction is more specific than the ones given above and therefore I will once more formulate them here. *Actus reus*, justifications and excuses are all ways of defending. A good defense makes the judge not convict the defendant. The first way to defend yourself is by lack of *actus reus*. An *actus reus* can simply be seen as the action or conduct of which the defendant is accused. With this goes the term *mens rea*, guilty mind. According to the law<sup>3</sup>, a person is guilty if he not only planned to commit the conduct but also actually committed it. So if a person did not commit the *actus reus*, even he had plans to, *mens rea*, he is right to say: ‘I did nothing wrong’. Once the *actus reus* has taken place another way of defense is in order; justification. Westen points out that there are some different opinions whether a distinction between excuse and justifications exists. However, most researchers agree with the point made above and Westen forms a good distinction between the two. Both by first way of defense as well as by justification the defendant is able to say: ‘I didn’t do anything wrong.’ This means even though he committed the *actus reus* this was not a bad thing due to the context it was done in. For example: When a policeman shoots a dangerous murderer, he actually does something good. Even though killing *prima facie* is a bad thing to do. What then is a defense by an excuse? Westen formulates this as every other way in which somebody can be defended, which is not a lack of *actus reus* or a justification. One important excuse is by lack of *mens rea*. A person who misinterprets a situation and thus presupposes he is acting in a proper way but is mistaking about this will be of no harm. However this leads to a major problem as is illustrated in an article of Montmarquet [14], that if the person acts right according to his own values, but these values are in conflict with the law, this will not count as a good defense. The person’s ‘mistake’ has to correspond with the law. Thus if the person was not mistaken his defense would be a justification. Again this is a counterfactual.

## 2.4 Summary

As stated in the beginning of this chapter the three different philosophies only amplify each other. The distinction between justifications and excuses made in ethics and language philosophy is extended in philosophy of law by integrating *actus reus* to it. To be regarded as an excuse, the committed conduct is still bad after the excuse is made but the observer no longer

---

<sup>3</sup>In American law.

holds the excuser responsible for the conduct. To accomplish this some restrictions have to be kept in mind. First of all, an excuse exists of two parts: the illocutionary part named  $F$  and the propositional part  $p$ .  $F$  is the force of making an excuse,  $p$  is the propositional part.  $p$  has to be the cause of the change in beliefs of the world by the observer. Due to this change of beliefs the excuser is no longer held responsible. Since an excuse is able to change something in the world it can be seen as an action. Another important restriction is that the conduct for which is excused has to be done unintentionally; however it has to be an action, which can be done intentionally. If the conduct is committed intentionally two options for excuses are offered, both of which make use of counterfactuals. First, when the excuser was forced to commit a certain conduct he will not be morally responsible if he did it only because he could not do otherwise. Second - and this originates from law philosophy - if a person is misinterpreting a certain situation but would have been justified for his conduct if the situation were as he interpreted it, he has an excuse.

## Chapter 3

# Psychology

With the use of philosophy theory, I have formulated an interesting but purely theoretical definition of excuses. With the use of psychology I will check if this theoretical idea matches reality. Psychology uses a different approach in probing excuses. Hypotheses formulated in psychology are tested in the real world instead of merely meditating arguments, which is done in philosophy. Therefore, by taking into account findings from psychological research my definition of excuses will be enriched. Looking at psychological research it becomes apparent that until the late '80 most research on excuses focused on when and where excuses occur. However, since then researchers have changed their focus. The focus now lays on the question if excuses are positive at all for the maker. This is quite a different strategy compared to the strategy philosophers follow. To give the full picture I will first review research, which has been done before the change. After that I will discuss the more recent literature.

### 3.1 What is an excuse?

For this part of my thesis I have used *Excuse: Masquerades in Search of Grace*. Snyder, Higgins and Stucky [22]. The book is a review of what has been written about excuses in psychology until 1983. Many more recent articles cite this book as giving an adequate definition of excuses.

According to Snyder et al, people are sometimes unhappy with themselves and make excuses to no longer be unhappy. Excuses occur when we feel we, not wanted, no longer look like a 'perfect' person. Excuses are used to repair the person's image. People want to be perfect or at least they want to have a positive self-image. Thus the purpose of an excuse is to keep ones self-image positive. It is made right after, at the moment or even sometimes before something bad happens. As stated in the introduction of this thesis I will focus on excuses made afterwards. Because a person's self-image is not threatened by good conduct by the person, it is only necessary to excuse for bad conduct. Due to this fact three essential conditions for an excuse occur. First, some action has to be done by the excuser. Second, this action is not good and the excuser does not want to be held responsible for it. Third, there has to be some sort of observer. People do not want other people to think negatively about them. Based on this Snyder, Higgins and Stucky [22, p. 4] give the following definition: *Excuses are explanations or actions that lessen the negative implications of an actor's performance, thereby maintaining a positive image for oneself and others.* They make a division into three kinds of possible excuses. These are named as follows: First, I did not do it. Second, it was not so bad. Third, yes, but... As those names already show, only one of these categories will be seen as an excuse according to the philosophical definition. In the first excuse, the excuser did not commit the conduct and therefore did not do anything wrong. The second kind of excuse can be seen as a justification and therefore it is not an excuse. Only the last one remains, the excuser admits he committed

the conduct, however he still thinks he is not liable. Since I only want to discuss excuses made in this last category I only describe the kinds of excuses which can be made in this category: First, *I couldn't help it*, means the only reason the excuser committed the conduct was because of a force he was not in control of. Second, *I did not mean to*. In this category all excuses fall which argue that the conduct was an accident. It was not the excuser's intention to commit the conduct. Third, *It was not really me*. Here the excuser dissociates himself as a person from his action. He states that if he had been himself he would have never committed the conduct. In short the three categories of excuses are: coercion, mistake and dissociation of the person.

As said before people prefer a positive self-image. This is one of the main reasons for making an excuse. A self-image can be defined as: *An actor's mental picture of his or her performance in a particular situation before audience*. [22, p. 28]. The self-image is thus the way a person thinks he appears to others due to his behavior. A person prefers a certain self-image. Therefore, if he thinks he committed conduct, which threatens this self-image, it will be affected in a negative way. Even though a person himself creates his optimal self-image it is influenced both by how the external audience and the internal audience react on the person. People do not want to get negative feedback because it will affect their self-image in a negative way. This proves that the external audience influences the self-image of a person. Since children use feedback to form their own opinions they develop an internal audience. This internal audience can give negative feedback as well. Thus this internal audience too has influence on the self-image.

However in order to get a good intuition of what an excuse is, it is not enough to know that people make excuses to maintain a positive self-image. It is important to know what kind of actions are threats to the positive self-image. The link between an action and a person is responsibility. Only when a person is responsible for an action, the action can affect the person's self-image. This can both be positive and negative. However, the fact that a person can be held responsible for an action is not enough. He actually has to be held responsible for the action as well. Someone only be held responsible for a bad performance by the observer when the following three points are met: First, the person is not performing under duress. If the actor is performing under duress he has no choice so the observer thinks he would have acted in the same way if he himself were in that situation. Secondly, the observer sees the actor as different, either psychologically or behaviorally from himself. Otherwise the observer can identify himself with the actor and will therefore think that everybody shall act that way in this situation. Lastly, the observer cannot be involved in a psychological way in the situation, since this can also lead to an identification with the actor. If one of those three points does not hold, the observer will not evaluate the conduct as bad and therefore an excuse will not be needed.

If, on the other hand, the three points do hold, the excuser should make an excuse in order to no longer be held responsible. Two things are important to keep in mind for making a proper excuse. First, excuses which are only self-reported can be biased by the actors' point of view and therefore do not correspond to the observers knowledge. Excuses which are (partly) physically manifested are therefore much more effective. Second, the more negative the action, the stronger the excuse should be. It is enough to say: 'I forgot what time it was.' when you arrive late at a party. But when a doctor uses the same excuse for explaining why he did not perform surgery on a patient who has died because he did not receive the surgery he needed, the excuse is not strong enough. Therefore it is important that the excuser evaluates the conduct correctly so that he knows how strong his excuse should be.

In short, in psychology an excuse is seen as a way to keep a persons self-image positive. An excuse is needed when this self-image is threatened. This can only happen when the person commits a certain bad conduct that he is held responsible for. By making an excuse the person can distinct himself from the conduct and thereby keep his self-image positive. it is important to notice that the way in which the word excuse is used in psychology slightly differs from the

use in philosophy. I will further discuss this in chapter 5, Integration.

## 3.2 How do excuses work?

Only a few years after the publication of *Excuse: Masquerades in Search of Grace* the interest of the researchers shifted a bit. As it was clear now what an excuse is and when it appears, researchers became interested in when excuses do keep the self-image positive, and if so, how excuses do this. Snyder and Higgins [23] describe different aspects of when excuses appear to have a positive outcome on issues such as anxiety, health, depression, self-esteem and performance. In these cases excuses do work. To make an excuse work five steps have to be followed. First of all, an excuse has to originate from some negative outcome of certain conduct of the excuser. In the second step the excuser dissociates himself from the conduct. Consequently, in the third step the excuser no longer focuses on himself. During the fourth step the excuser experiences the positive effects of making an excuse, followed by step five, which is maintaining a positive self-image by the excuser.

As already mentioned, responsibility is an important concept for making an excuse. To find out more specifically how making an excuse works, it is interesting to probe responsibility a bit further. Making an excuse is the attempt to no longer be held responsible for certain conducts. This is what Schlenker et al noticed. It is important to keep in mind that responsibility is often caused by obligations made by some other person [16]. Schlenker et al made the so-called *Triangle Model of Responsibility* [15]. This theory is based on the idea that responsibility is an essential element to hold people accountable for a conduct. Being accountable is the justification for others to judge the person who committed the conduct in a negative way. A person who committed a bad conduct should first be evaluated before he can be judged. Three key elements should be evaluated:

1. *Prescriptions*, these let a person know how to behave in certain situations.
2. *Event*, this is what is evaluated, thus the bad conduct that is committed.
3. *Identity*, this refers to the part of the actors' personality which has to do with the evaluated conduct.

All three elements are interlinked. As shown in figure 3.1 this looks like a triangle. The links together represent responsibility. The stronger the links the more responsible a person is for the conduct. The link between *prescriptions* and *event* shows how clear the prescription was. If the prescription of a certain request or obligation was unclear to a person, he is less accountable for the conduct. The *prescriptions-identity* link shows to which point the person is bound to the prescriptions, thus the extent to which the prescriptions are internal. If this is high, the person is more accountable. The last link is between *event* and *identity* and is about how much control the person has over the outcomes of the action. The higher the control the higher the accountability. However, not only the links are important to know in order to what extent a person is accountable.

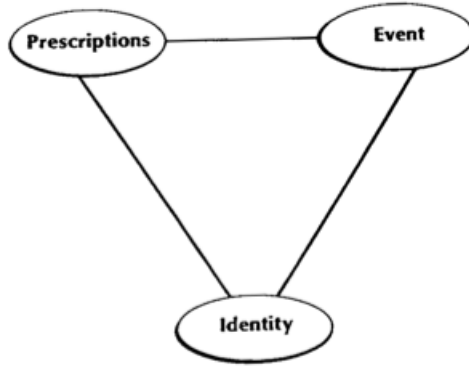


Figure 3.1: The responsibility triangle [15].

It is also important to consider the so-called *potency* of the elements. If an element is more important to the judger the accountability is higher. The combination of the three links between the elements is seen as responsibility. The impact an event can have on a person depends on two things. First the potency of the elements and second, the degree of responsibility. If those two increase, the person gets a bad evaluation of the conduct, which can lead to a negative self-image. Since this is what excuses try to prevent, weakening the links is what happens when somebody is making an excuse. This however is not the only way in which the elements and its links can be used. The links also play a role before an action. If the links are strong the person feels a high responsibility to succeed in performing the action. The person is more ‘willing’ to fulfill the request or obligation than if the links are less strong.

### 3.3 Negative site of excuse

Until now it seemed as if making an excuse is by definition a good thing to do. It helps people to maintain a positive self-image because it keeps them away from bad evaluations by others. Many books and articles in psychology recommend people who are failing at some point to make excuses to keep their self-image positive [16]. However, in the last decade the focus of research in psychology changed and psychologists are no longer only focused on the positive side of excuses. They also use the outcome of research on negative aspects to get a better insight in what a proper excuse should be. Sheldon and Schachtman [20] found a negative association between future commitment and excuse making. They used the triangle model of responsibility and compared this with another theory named; *Self-determination theory* (SDT). This theory partly overlaps with the triangle theory. It only differs from the triangle model in the nature of the link between identity and prescription. SDT is very strict on this point. It definitely has to be the persons own values, which make the link strong and not the values of the obligator/observer. The outcomes of the studies, which are reviewed in the article of Sheldon and Schachtman, match with the hypotheses made in SDT and the triangle model. If a person does not make an excuse, there is a bigger chance he will behave differently the next time. So you could say that making excuses saves somebody’s positive self-image but the person will learn less from his actions. Therefore the chance that he will again commit the conduct is higher if he excuses for the conduct.

Especially in social groups, making a lot of excuses can lead to problems. Social groups are healthy if the people in the group are reliable. If people no longer act the way they are expected to act, this leads to problems. Excuses give people the ability to undermine reliability. If a person makes excuses too often he will no longer be predictable for the other members in the

social group. This is why making excuses is sometimes seen as improper. Excuses can therefore be seen as a threat to the character of the excuser. With character Schlenkner means that a person behaves as is expected from a healthy person. To make a good excuse that does not have the disadvantage of showing a bad character some conditions have to be fulfilled: First, the excuse has to be credible. The excuse has to correspond to the knowledge of the observer and the excuser himself. Of course it helps, as stated before, if there is physical evidence. Secondly, the excuse must only dissociate the excuser from the part of which it is correct to do so. Lastly, the excuser must maintain good will. As excuser you have to be sympathized with by the observers. An excuse made by a sympathetic person will be accepted more easily than an excuse made by an unsympathetic person. These three conditions make clear that lying to have a good excuse does not work well. Therefore it is necessary, in order to be able to successfully make an excuse for a certain conduct, that the conduct was in some way not the responsibility of the excuser.

However this solution assumes that if excuses are not misused there is no problem to the health of the group. I doubt that. Even though the excuse is sincere and justly it still is a form of weakness. Not doing anything wrong seems better to me than doing excusable bad things. This however brings up a different discussion, which cannot be held in this thesis, due to limited space and time.

### 3.4 Summary

The definition of what an excuse is, is broader in psychology than in philosophy. Both justifications and excuses after which the conduct is still bad are seen as excuses. The articles I used are all about the second kind of excuses. Three kinds of conduct are discussed for which can be excused: coercion, mistaking and acting in a way, which does not fit the personality. The purpose of making an excuse is to maintain a positive self-image. This image is threatened if a person is held responsible for bad conduct and therefore bad conduct is necessary to make an excuse. The observer can only hold the excuser responsible if the following points are met: There is no duress, the observer sees the excuser as different from himself and the observer is not emotionally or psychologically involved in the situation. If the excuser is held responsible and made an excuse the following steps occur: An excuse dissociates the excuser from the committed bad conduct. After that, the excuser no longer focuses on himself and the excuser experiences the positive effects of the excuse. Finally the self-image is repaired. Even though, until now, excuses seemed quite positive for people, there are some negative effects on them as well. First, if a person makes an excuse for his mistake, he learns less of his mistakes than if he does not make an excuse and therefore is held responsible for the mistake. Second, in social groups, people who often make mistakes are less reliable. As such it is necessary to keep the following aspects in mind to minimize the negative side effects: the excuse has to be reliable, credible, and only dissociates the excuser from the correct part. Therefore a physically manifested excuse is more effective. In addition the excuser has to maintain goodwill from the observer.

# Chapter 4

## Logics

It seems that not a lot has been written about the concept of excuses in logic. There are only a few articles written about excuses in general. However a lot has been written about different concepts on their own which are, as shown in psychology and philosophy, aspects of an excuse. This chapter will therefore have a different approach compared to the other disciplinary chapters. I will first discuss some articles about excuses written by logicians and thereafter I will use the restrictions on excuses found in philosophy and psychology to be able to write about excuses from the discipline Logic. However, I will not yet integrate the way in which those concepts are used or described in philosophy or psychology. I will simply take over the definition of the different concepts as stated in logic. By doing this I will minimize the bias of the other disciplines in logic. Therefore, after this chapter, three almost individual definitions of excuses will be formulated which can be integrated into one overarching definition.

### 4.1 Concepts

In logic it has already been noticed that more than one agent is needed for making an excuse. Therefore multi-agent-models are necessary [4]. Boella makes use of Austin's definition of excuse and his distinction between an excuse and a justification. According to Boella et. al., an agent behaves as he is obliged to. When he acts in a different way, he will be held responsible and punished for his actions unless he makes an excuse for his behavior. There are five categories of excuses: epistemic excuses, power-based excuses, norm-based excuses, counts as-based excuses and social-based excuses. The first category is about excuses on the absence of knowledge about the obligation you violated. In the second category, power-based excuses, all excuses are collected, which show that the agent did not have the right skills to fulfill the obligation. The third category contains norm-based excuses, which are used when an agent prefers to do something else because that thing was more important. The counts as-based excuses form the fourth category, and is a sub-category of the epistemic category. An excuse belongs to this category if it can be said that the violation did not occur because it did not count as one. However, these two previous categories seem to me more like justifications than excuses. If a violation did not count as one, there is no longer bad conduct. This also happens in a norm-based excuse. If the observer agrees with the excuser that the other thing was more important he will not value the conduct of the excuser as bad. Therefore it is a justification. The last class is of social-based excuses, probably one of the most used excuses in this class is: "Everybody does it".

In an other article Broersen [8] analyzes the concepts of *actus reus* and the levels of *mens rea*. He claims that the levels of culpability correspond to the levels of excusability. He states that certain concepts have influence on this. These concepts are: knowingly doing, deliberateness



and obligation-to-do. As became clear in the other chapters much more concepts are related to excuses.

The concepts, which are aspects of excuses, can be divided into three categories. Some concepts are restrictions for the conduct, which is excused for while some are restrictions for the excuse itself and the last category will hold restrictions for the audience. I shall discuss every concept on its own. If a concept occurs in psychology as well as in philosophy I will not discuss it twice. I will only discuss the concept as it is defined in logic. In the next chapter I will integrate the vision of the different disciplines on each concept to one overarching definition.

## 4.2 Restrictions for conduct

All models that represent a certain restriction for the conduct of excuse are a form of STIT logic. Therefore I will first shortly introduce this kind of logic before describing the restrictions. STIT stand for ‘sees to it that’. So the human-language translation for  $[A \textit{ stit}: \varphi]$  is: “A (an agent) sees to it that  $\varphi$ ” [12]. The articles I will use have a different notation, which is more in line with the standard modal logic notation. Therefore I will use the following notation [8]:

$$[A \textit{ stit}]\varphi$$

The central operators in STIT logic represent action. I will make use of a *stit* logic in which the effect of the action occurs in the next state: XSTIT. This XSTIT logic is defined in an article of Broersen [8]. Here is a visualization of a simple two-agent XSTIT frame:

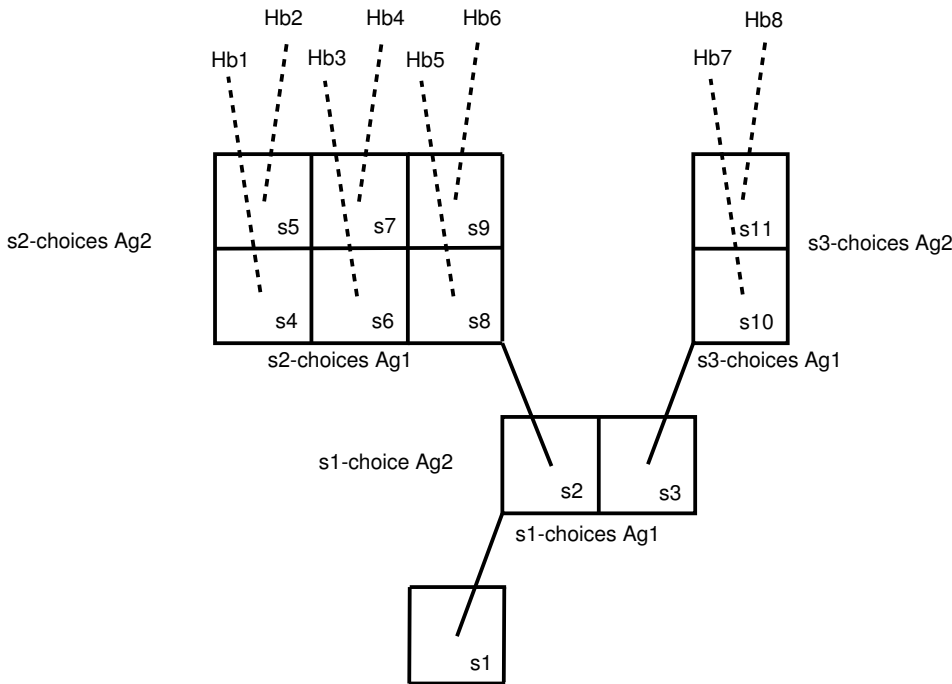


Figure 4.1: Visualization of a simple two agent XSTIT frame.

Choices are actions. By doing a certain action you make a choice. The small squares represent static states. Ag1 his actions are represented as columns and ag2 his action are represented as rows. Thus from s1 only ag1 has a choice. If s3 is the outcome of his choice only ag2 can decide if s11 or s10 will be the next outcome. The line that can be pulled between those states is the history. Since on XSTIT the outcome of the choice appears in the next state

this line is infinite. Hb stands for the bundled line of different histories. Since from most states it is possible to make choices different histories are possible.

For some restrictions it will be useful to know what a single agent ‘view’ on a multi-agent XSTIT frame looks like.

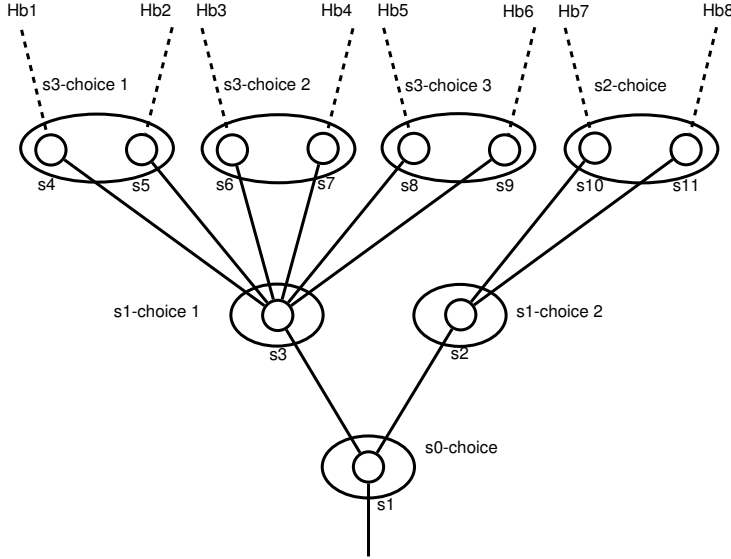


Figure 4.2: Visualization of a simple one agent ‘view’ on a two agent XSTIT frame.

Static states are pictured as little circles. The ellipses visualize the different choices of which Ag1 is in charge of. The other lines and names depict the same as they do in figure 4.1.

## Obligation-to-do

The restriction I will start with is that the conduct has to be bad. A conduct is bad when an agent fails some sort of moral obligation. The obligation-to-do model [7] shows exactly when an obligation is violated and thereby when bad conduct has occurred. Thus an excuse has to be made if and only if a violation occurs. The definition of an obligation is clear: *An agent is obliged to do something if and only if by not performing the obliged action, it performs a violation.* To model this, XSTIT logic is used. The definition of the obligation-to-do model is as follows:

$$O[a \text{ xstit}] \varphi \equiv_{def} \Box(\neg[a \text{ xstit}] \varphi \rightarrow [a \text{ xstit}] V)$$

$\neg[a \text{ xstit}] \varphi$  means that  $A$  ‘allows’ a choice for which  $\neg\varphi$  is a possible outcome. Due to the implication that if such a thing is allowed it will definitely lead to a violation, this obligation is a personal obligation. If some other agent by coincidence acts in a way such that  $\varphi$ , this still leads to a violation for agent  $a$ . In the article of Broersen a stronger definition is discussed in which the agent is aware of actually performing the obligation and otherwise it will count as a violation. This however is not part of the restriction for a conduct of which can be excused for. It does not really matter for what reason a person holds completely to an obligation. Thought the last definition stated in the article is interesting. This definition holds that the agent is aware of the obligation and refrains from meeting the obligation. The extra operator  $K_a\varphi$  stands for knowledge of individual agents  $a$ .

$$KOK[a \text{ xstit}] \varphi \equiv_{def} \Box(\neg K_a[a \text{ xstit}] \varphi \rightarrow K_a[a \text{ xstit}] V)$$

In the next chapter I will discuss this definition further. Because I do not want the agent to necessarily be aware that he is refraining from the obligation, but in some situations it is needed that the agent knows about the obligation, this definition is not perfect for excuses.

## Intentionality

A second important restriction to the conduct can be found in the concept of intentionality. It should be possible for the agent to commit the conduct intentionally, but in the case of which the agent makes the excuse, the agent committed the conduct unintentionally. The conduct should ‘happen’ to the excuser. According to Bratman [6] there is a difference between intentionally doing something and intending to do something. Since I am interested in the first, I will use the logic that prescribes intentionally doing, studied by Broersen [9]. In this thesis intention is defined as a mode of acting and therefore XSTIT logic is eligible to use. For intentionality it is better to use a one-agent ‘view’ such as figure 4.2.

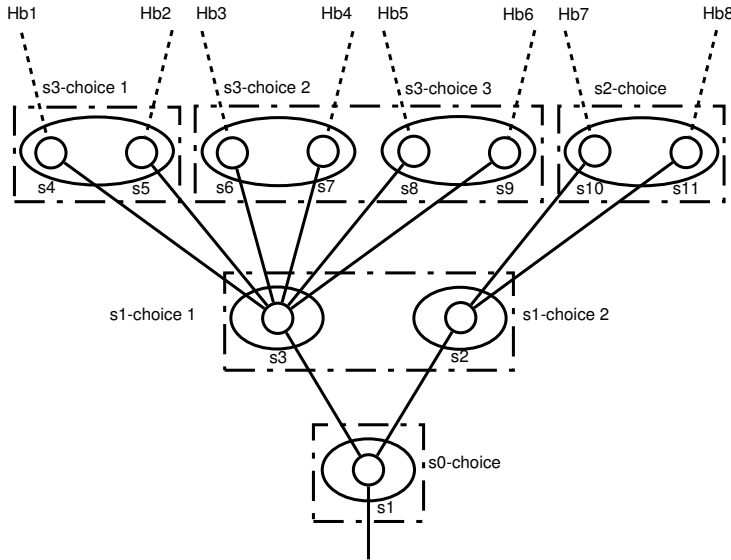


Figure 4.3: Knowingly doing in a K-extended XSTIT frame.

Besides objectively making a certain choice, which is modeled by basic XSTIT logic, and intentionally making a choice there is another important way of choosing. This is what an agent knows or believes he is choosing. As seen with the first restriction, XSTIT logic can be extended with an operator  $K_a\varphi$  for knowledge. For the intentional level the operator  $I_a\varphi$  is used. Which means: ‘agent  $a$  intends  $\varphi$ ’.  $K_a[a \text{ xstit}]\varphi$  models knowingly doing, which models an action that the agent is aware of doing, and  $I_a[a \text{ xstit}]\varphi$ , which models an intentional action.

Knowingly doing is depicted in figure 4.3. The dotted lines show which choices can be made knowingly. Thus although in  $s_3$  Ag1 can objectively choose between three choices/actions he can only choose knowingly between  $s_3$ -choice 1 and the set of  $s_3$ -choice 2 and  $s_3$ -choice 3. Ag1 cannot knowingly differentiate between  $s_3$ -choice 2 and  $s_3$ -choice 3. Broersen [9] defines more properties in his article. It is not possible or necessary to describe them all here, since they are not of explicitly necessary in order to form a definition of excuses.

Intentional action is depicted in figure 4.4. The dotted ellipses depict the sets of choices between which can be chosen intentionally. The other lines and symbols are as stated before. So if from  $s_1$  in figure 4.4 the next objective state is  $s_3$  ag1 intended to do what is held in  $s_4$ ,

s3 and s2 and ag1 knows to be doing the action depicted by the dotted rectangle around s3 and s4. Thus in figure 4.4 ag1 had two possible intentional choices.

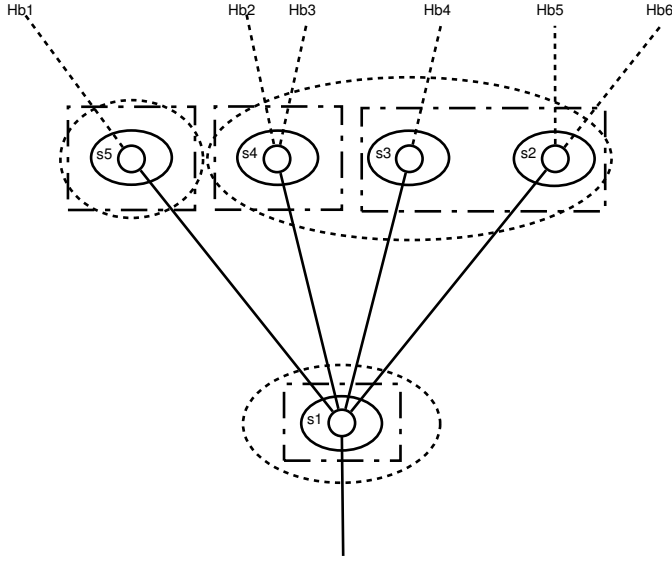


Figure 4.4: Knowingly doing and intentionally in a KI-extended XSTIT frame.

An important restriction on intention for this thesis is that intentionally doing implies knowingly doing thus:  $I_a[a \text{ xstit}]\varphi \rightarrow K_a[a \text{ xstit}]\varphi$ . This brings the following question: what is the difference between knowingly doing and intentionally doing? An important aspect for this question is deliberativeness. In normal XSTIT logic a deliberative action is defined as follows:

$$[a \text{ dxstit}]\varphi \equiv_{def} [a \text{ xstit}]\varphi \wedge \Diamond X\neg\varphi$$

$X\varphi$  means that the next state ensures  $\varphi$ . Thus, to do things deliberative according to XSTIT logic an agent has to have another choice. From this can be deduced that to do something intentional the agent must be able to knowingly refrain from what he is doing. Therefore the following definition is the final definition of a model of intentional action.

$$[a \text{ xint}]\varphi \equiv_{def} I_a[a \text{ xstit}]\varphi \wedge \Diamond\neg K_a[a \text{ xstit}]\varphi$$

This definition brings a solution to the problem that intentions should be closed under knowledge. Broersen [9] argues that it is not desirable that someone who intends to go to the dentist and knows going to the dentist causes pain intends to have pain. Therefore the following axiom (SE) is not wanted:

$$I_a[a \text{ xstit}]\varphi \wedge \Box(K_a[a \text{ xstit}]\varphi \rightarrow K_a[a \text{ xstit}]\psi) \rightarrow I_a[a \text{ xstit}]\psi$$

This axiom however is derived from the base system as becomes clear in figure 4.4. Everything, which is in a dotted rectangle, is also in the same dotted ellipse. This means SE is too strong. The only thing, which is certain when going to the dentist, is risking pain. According to Broersen the solution to the problem is the following: *The side effect problem is no problem if we look at an intentional act as the outcome of a process of deliberation.* [9, p. 521]. According to this, the intended action is the action, which is done deliberately. All the side effects caused by this action do not occur deliberately, and therefore are not intentional. Broersen's definition of  $[a \text{ xint}]\varphi$  holds to these restrictions.

One problem with the final definition that is not yet solved is the problem mentioned in philosophy. The restriction of  $\diamond\neg K_a[a \text{ xstit}]\varphi$  no longer applies if the chosen action does not change when the agent gets other options. Since there is no solution found in logics, I will further discuss this in the next chapter.

## Mistake

The third restriction is about mistakes and originates from philosophy. Agent  $a$  can be excused if the committed conduct would have been justified if the situation was really like agent  $a$  interpreted it. So agent  $a$  made a mistake by interpreting the situation in the wrong way. Mostly this happens when an other agent's choices are involved, agent  $a$  misinterprets the intentions of agent  $b$ . An extension of STIT logic that takes this into account is attempt [10], which is called XSTIT<sup>p</sup>. In this logic probability is taken into account. An action can be called an attempt if an agent chooses the action, which has the highest probability for the wanted outcome. This does not mean the probability has to be high. This idea can be used for modeling a mistake. Probability has to do with mistakes in a slightly different way. Agent  $a$  interprets a situation, maybe not explicitly, as is most plausible for him in the current context. Since the operator for attempt is not yet defined in the literature I cannot make use of the ideas proposed above.

## Responsibility

The last restriction is about responsibility. Since there is not much literature about this in logic I will come back to responsibility in the next chapter in which I will formulate an operator and a definition.

## 4.3 Restrictions for excuse

Philosophy showed that an excuse can be modeled as  $F(p)$  in which  $p$  is the propositional part. An excuse is proper if  $p$  causes the fact that the excuser is excused. Besides this  $p$  may not lead to any contradiction which of cause leads to  $\perp$ . This means  $p$  has to be coherent with other facts known by the observer and may not lead to a falsum. In the next chapter I will start integrating these restrictions with the other restrictions.

## 4.4 Restrictions for the observer

According to philosophy, there are not many restrictions for the observer. The observer does not have many restrictions according to philosophy. In psychology there are some more restrictions, which the observer must meet. However none of these restrictions are discussed in logics. Therefore the translation into logic will happen as much as possible in the next chapter.

## 4.5 Summary

As seen, most restrictions are on the conduct itself. These are also the restrictions about which most is written in logic. For all restrictions about which is written in logic, XSTIT logic is used. This multi-agent-logic is able to talk about the next state. The found operators have to be revised or extended to work better for excuse. For some restrictions, especially on the excuse itself and for the observer, new operators have to be proposed. This will be done in the next chapter.

## Chapter 5

# Integration

In this chapter I will integrate the different definitions of the restrictions or concepts of an excuse as found in philosophy, psychology or logic into one overarching definition of excuse and its restrictions. I will start by giving an overview of the different perspectives of the different disciplines to excuses in general. Next, I will discuss every restriction made, by at least one of the disciplines, and compare this to perspectives of other disciplines to this restriction. Once the final definition of the concept is clear, this will be visualized with a logical model of that concept. All these concepts will be integrated in one logical model, which visualizes the definition of excuses. To be able to make the model I sometimes have to propose new ideas for a logical representation of restrictions on excuses. Since this is a bachelor thesis I will make use of semi-formal logic. I am not able to formally prove my ideas. So these ideas have to be seen as just an option. I sometimes will even make use of other logics than STIT logic, since there is not - maybe not yet - the possibility to model the things I want to model in STIT logic. For some restrictions it will even be too difficult to propose a semi formal logic. If this is the case, I will not propose any logic and keep to human language.

### 5.1 Perspectives

In reading and writing for this thesis it occurred to me that psychology and philosophy look very differently at excuses. Since I used concepts found in those disciplines for the chapter about logic, I cannot say much about the perspective on excuses in logic in general. The few logical articles about excuses in general all have a philosophical background. However, something interesting happens in the article of Boella. I will shortly discuss this at the end of this section.

In philosophy the idea of what an excuse can be is much more narrowed down than in psychology. Most articles in philosophy explain at the start of the article what they mean with the word excuse and exclude the possibility of an exception to this definition. The definition of an excuse in philosophy is normative. In psychology the definition is much more descriptive. Psychologists search for exceptions and interpretations of excuses in daily live by doing research. I suppose this is why the definition of what counts as an excuse is much broader in psychology. Even though psychologists do agree with philosophers that there is a distinction between *actus reus*, justifications and excuses as defined in philosophy, psychologists state that these are all forms of excuses. Therefore, sometimes it is hard to notice if an article in psychology is about excuses or justifications. I am interested in the form of excusing in which the conduct stays bad but the committer is no longer responsible. Therefore, I only used assumptions and statements made in psychology on this form of excuses. In some articles of psychology even apologizing appears to be an excuse. With apologizing it is the same as with justifications. Psychologists do make the differentiation, however sometimes the four get mixed up. Probably because it

often happens that people use these four different ways of making their self-image more positive at the same time.

This reason directly leads to the next distinction between the two perspectives. Both disciplines have restrictions for the conduct itself, the excuse and the observer. However philosophy concentrates much more on the conduct and ascribes most restrictions to the part of what is needed for an excuse. On the contrary, psychology focuses much more on people, on the excuser and how he should make an excuse, and the observer. It is interesting to see that most restrictions appear in philosophy as well as in psychology. However in philosophy these restrictions are mostly described via the conduct and in psychology via the audience or via the excuse. Since, in my opinion, many of the restrictions on the excuses made by psychology are actually restrictions on the conduct, I have most of the time chosen to keep to the division as made by philosophy. This is defended in the next section.

In the chapter on logic I mentioned the article by Boella in which five ways of excuses are stated. These five ways of excusing can be seen as restrictions on the excuse. However, I think it is better to connect these restrictions to the conduct and to the observer instead of on the excuse itself. The conduct prescribes which facts should be mentioned in the excuse and the observer gives the reason for the agent to excuse: he has to accuse the excuser of bad conduct. So the context creates the excuse. Therefore I will discuss those three ways of excusing with their restrictions. The other two restrictions are, as stated in the chapter of logic, forms of justifications.

## 5.2 Restrictions

Since all three disciplines agree on the fact that an observer has to be involved, I will make use of multi-agent models as is already done in logic. According to philosophy an excuse is only a proper excuse if the observer accepts it as such. This means there has to be an observer who does not agree with the conduct. In psychology it has been stated explicitly that an observer is needed before an excuse makes sense. According to psychology, the aim of making an excuse is to keep the self-image positive. The self-image becomes negative if a person is held responsible by an observer for the bad conduct committed. Thus at least two agents are needed in a model of making excuses.

The model, is not only a multi-agent model, it also is an action-model. This is also already stated in logic and agrees with the statements in philosophy and psychology. As seen above, the reason for making an excuse is to no longer be held responsible for a certain conduct. To achieve this, the observer's beliefs of the world need to be changed. Both in psychology and in philosophy restrictions are found on how this is done. According to philosophy an excuse is only proper if the change of view of the world is caused by the assertion ( $p$ ) in the excuse. For psychologists it is slightly less clear that the excuse has to be the cause of the change but they do give all kinds of restrictions to the excuse and state that to make a good excuse the links in the triangle model, which visualizes responsibility, have to be weakened. Thus both psychology and philosophy would agree on the idea that by making a proper excuse a person *does* something. Therefore an excuse can be seen as a certain action.

Because of this it seems plausible that the decision made in logics to make use of XSTIT logic works as both for the definition of excuses in psychology as in philosophy. For some parts of the restrictions it is preferred to use another kind of logic. This will be explained with the particular restriction.

The restrictions have a chronological order. Some restrictions only occur if some other restrictions are fulfilled. In the following timetable the different steps are stated. These steps

originate from both philosophy and psychology. Next, I will discuss each step and its restrictions.

t1	t2	t3	t4	t5
Action done by agent $a$	Observer holds agent $a$ responsible	Agent $a$ excuses for action	Observer's view of world changes	Agent $a$ is no longer liable for action

In t1 the action is done. Thus all restrictions to the actions have to be checked to know if, according to t1, an excuse is in place. The first and most important restriction, supported by psychology and philosophy, is that the action committed by agent  $a$  has to be bad. As discussed above, psychology sees this as one of the classes of excuses and I decided to only probe that class. Philosophy on the other hand states that an excuse can only occur when an agent committed a bad conduct. In the chapter on logic I discussed that doing something bad can be seen as not following the rules of a certain obligation. This leads to a violation of the obligation. Thus in order to make an excuse a violation has to occur. The XSTIT logic obligation-to-do helps to visualize this restriction of an excuse:  $B[a \text{ xstit}] \varphi \equiv_{def} [a \text{ xstit}] \varphi \rightarrow [a \text{ xstit}] V_a$ .  $\varphi$  visualizes the result of the bad conduct.  $V_a$  symbolizes the violation that occurs for a particular agent, in this case agent  $a$ , if the bad conduct is committed. For this definition to be true it is not necessary that the agent himself commits the bad conduct. However this is a restriction both in philosophy and in psychology. Therefore the definition needs a simple correction. If it is not the agent himself who commits the conduct, it will not lead to a violation by that agent. It is not possible to excuse for some other agents' actions. My final proposal for agent  $a$  committed bad conduct by seeing to it that  $\varphi$  is:

$$BC[a \text{ xstit}] \varphi \cong ([a \text{ xstit}] \varphi \wedge \neg \Box X \varphi \wedge \neg [b \text{ xstit}] \varphi) \rightarrow [a \text{ xstit}] V_a \text{ for: } a \neq b$$

In addition to this, philosophy states that it must be able to commit the conduct intentionally. This is in line with the idea of psychology that being held responsible for conduct is a reason to excuse. In the triangle model of responsibility it becomes clear that this has much to do with intentionality. The link between *event* and *identity* is about the amount of control the person has on the outcomes of the action. If this is high, the accountability is high as well. This means that a person has to be able to be in control over an action to be held responsible. If a person acts intentionally, he wants to act that way and he also knows he is acting that way. He therefore himself causes the action and is in control over the action. It might look as if knowingly doing is important here as well, but it is not. In my opinion everything that can be done knowingly can be done intentionally. Because the only difference between the two, is that with intentionality it is necessary to want the action as well. Even though for some action it will be hard to imagine a person would want to act that way it is still possible. In logic a definition of intentionality is given.  $[a \text{ xint}] \varphi$ . To go from t1 to t2 the action in t1 has to have the possibility to make this definition true. Thus I propose that it is possible that agent  $a$  intentionally sees to it that  $\varphi$  should be symbolized as follows:

$$\diamond [a \text{ xint}] \varphi$$

It is not possible to excuse for all actions, which holds for the above-described restriction, in the same way. Therefore the next restriction will divide the action up in different categories of ways of excusing. A possible kind of action for which Agent  $a$  can excuse is: if he only committed the conduct because he was forced to do so. In psychology being forced to do is seen as a way of excusing. But on the other hand, it is also stated as a restriction of the observer. An observer is only able to hold someone responsible for a certain conduct if that person did not



act under duress. This means that according to psychology a person who is acting under duress may not be held responsible and therefore t2 will fail. Since the observer is not able to know in every situation if a person acted under duress or not, it may be necessary to explain this by way of an excuse. Thus, to make a sincere excuse by stating that agent  $a$  was forced, the action must be forced. Therefore it seems more accurate to make this a restriction of action. In logic a restriction on this is found as well in the power-based excuses, which are excuses in which it is stated that the excuser did not have the right skills. If an agent is obliged to act in a certain way, but he is not able to act like that since he does not have the right skills, he is forced to act differently. Therefore this kind of excuse named in logics belongs to this restriction.

In the chapter on logic I explained the definition of deliberative XSTIT logic:  $[a \text{ dxint}] \varphi \equiv_{def} [a \text{ xstit}] \varphi \wedge \Diamond X \neg \varphi$ . Coercion can be symbolized as:  $\neg[a \text{ dxint}] \varphi$ . Since this only demands that no other option is possible it is not enough for this restriction. The part of the restriction in which it is excluded that the agent would have acted in the same way if he had other options has to be integrated. I propose to symbolize this as follows:  $\Diamond X \neg \varphi \rightarrow [a \text{ xstit}] \neg \varphi$ . However this part is a counterfactual, since it is in a hypothetical situation. Until now nothing is written about this in STIT logic and it is not possible to integrate this into one formula. Therefore I propose to introduce a new operator, which symbolizes the swift to another world at the same moment:  $\bigcirc$ . Thus agent  $a$  was forced to see to it that  $\varphi$  and if he was not forced he would have acted different is proposed to define as follows:

$$[a \text{ coxstit}] \varphi \cong \neg[a \text{ dxint}] \varphi \wedge \bigcirc(\Diamond X \neg \varphi \rightarrow [a \text{ xstit}] \neg \varphi).$$

Another category of actions, which can be excused for, is that of mistakes. This restriction originates from philosophy. The person who shoots an actor because he thought he was really threatened has an excuse for his conduct; because if the actor had been a real person the excuser would have had a justification. This can be associated with a link to the responsibility triangle found in psychology, which is the link between *prescription* and *event*. This link is about how clear the prescription was to the agent who commits the bad conduct. This of course is different from making a mistake, but it overlaps in some ways. By committing the conduct, the agent thought he was justified to do so. He misinterpreted the situation and with that the belonging prescription. The kind of excuse stated in the chapter on logic of absence of knowledge belongs to this restriction. Absence of knowledge means the agent does not have enough information to act in the right way and therefore makes a mistake. To show this in logic a model is needed which prescribes that the agent was mistaking and if the situation was as he interpreted the situation, he would have been justified and therefore no valuation would have occurred. The logic of attempt is linked to a logic of mistake and can therefore help to propose a new operator. With an attempt the best option is taken. I suppose that it is only possible to excuse for a mistake, if the agent's chose to interpret the situation as it was most likely to occur to him for good reasons. Thus with making a mistake the most likely interpretation has to be taken to make a good excuse. However, this brings with it a problem. It is subjective to say which interpretation is most likely, therefore in practice it is probably hard or even impossible to objectively state which interpretation is most likely. The question arises whether it would be better to say that the actual situation is the most likely situation. I leave this question to philosophers. For this thesis I keep to the idea that it is possible to know which interpretation is most likely to an agent. The symbol  $Bq$  means that  $q$  is the most likely interpretation [3]. The conduct, which he committed, does not lead to a violation if the situation would have been as he interpreted the situation. Again this is a counterfactual. I make use of hybrid logic to be able to state that in the counterfactual world  $q$  actually is the case. This together leads to the following idea for a definition for agent  $a$  mistakingly sees to it that  $\varphi$ :

$$[a \text{ mis}] \varphi \cong Bq \wedge \bigcirc (@q \rightarrow \neg(BC[a \text{ xstit}] \varphi))$$

The last category of actions is in some way related to the previous category. By making a mistake the agent believes he is justified in the way he is acting. He is not aware of the fact that he is doing something wrong. In the last category the agent is also not aware, at least not while he is preparing for the action, that he is doing something wrong. This category is about unknowingly doing. I already discussed intentionality. That has much to do with this category. As seen in the chapter on logic, knowingly doing is part of intentional doing. Both psychology and philosophy state that an action for which can be excused cannot be done intentionally. However this restriction is not strong enough. The example given in the chapter on logic makes this clear. Going to the dentist means going to be in pain. It is strange however to state that if an agent intentionally goes to the dentist he intends to be in pain (intentionality). However, he knows he will have pain (knowingly doing). If the example is not the dentist and having pain, but being at work on time and driving too fast it is not correct to say he has an excuse. In some situations this might lead to a justification, but never to an excuse. Thus, in order to be able to excuse it is necessary to not know about the side effect that is a bad conduct. The definition of this is very simple:

$$\neg K_a[a \text{ xstit}] \varphi$$

There is, however, a problem with this definition. Not knowing might be too soft. In some situations it is possible that the agent must not only not know that something can happen he must also not belief that it is going to happen, or a least that it is most likely to not happen. This however still is not enough in some situations. For example, in surgery it is not enough to be 55% sure that no instruments will be left in the body. It seems to me that it is up to the observer to decide whether the excuse made for the situation is acceptable. If the excuse made for the situation is excepted. I will discuss this point some later in this chapter. In order to be able to make an excuse it is enough to just not know as defined above.

It is time to go to the second step in the timetable, t2. The observer holds agent  $a$  responsible. The name of this step already is a restriction, which is both stated in philosophy and psychology. However, according to psychology before the observer is able to accuse agent  $a$  three condition have to be met for the relation between the observer and agent  $a$ . First, duress, I already discussed this restriction and shifted it to t1. Secondly, the excuser is seen as different from the audience and thirdly, which follows from the second, the observer must not be psychologically involved in the current situation, otherwise he could identify himself with agent  $a$  in the current situation. The last restriction leads to one of the kind of excuses statute in logic, everybody does it. If this is true the observer cannot hold the person responsible and therefore an excuse is not necessary. If it is not true it can not be used as an excuse since in an excuse the agent may not lie. This will be explained in t3. The restriction can be visualized as follows: the observer belongs in some way to a different class of people then agent  $a$ :

$$\text{observer} \notin A \wedge \text{agent } a \in A.$$

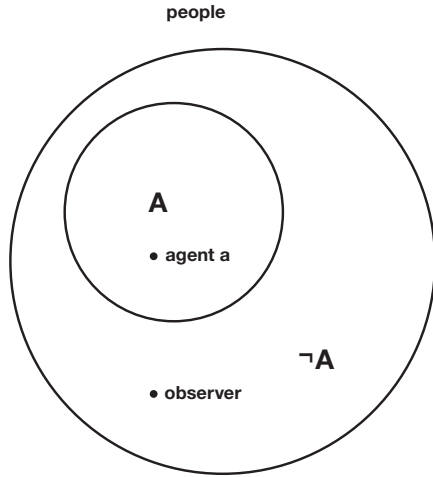


Figure 5.1:  $observer \notin A \wedge agent a \in A$

See figure 5.2. If this is true, the audience is able to hold agent  $a$  responsible. This means that the observer knows that agent  $a$  committed bad conduct. Thus I propose:

$$K_o BC[a \text{ xstit}] \varphi \text{ for } o \neq a$$

However, this is a bit strange since the observer must know agent  $a$  committed bad conduct in the past. In this formula it looks more like the observer knows that agent  $a$  is going to commit bad conduct. This cannot be visualized in STIT logic. I propose it should be like at moment  $t_2$  ( $q$ ) observer knows agents  $a$  committed bad conduct at  $t_1(r)$ :

$$@_q K_or$$

If these two restrictions hold, agent  $a$  has to excuse for the bad conduct,  $t_3$  in timetable. In the chapter on philosophy due to Austin and Searle an excuse is pictured as follows:  $F(p)$ . To make  $F$  successful it is necessary that agent  $a$  intentionally made the excuse. This can be associated with the condition found in psychology that agent  $a$  has to maintain goodwill. It has to be made clear to the observer that the excuser sincerely makes the excuse. The observer must get the idea that agent  $a$  would have acted in a different way if he could have foreseen the bad conduct. This is also a counterfactual, and I propose that it can be modeled as follows:  $[a \text{ xint}] \varepsilon$  in which  $\varepsilon$  denotes the situation that agent  $a$  excused (in his opinion) for the bad conduct.  $p$  on the other hand denotes the part of the excuse in which agent  $a$  states certain facts. Philosophy states that these facts have to be the cause of the change of the observers' view of the world. Since in psychology two restrictions are made to facts stated in the excuse to be accepted as such, it is clear that also in psychology it is believed that these facts are the cause of keeping the positive self-image. These restrictions are: First, the excuse has to be credible; it may not contradict with knowledge of the observer. Second, agent  $a$  has to stick to the facts, he may not lie. This means stating  $p$  may not lead to any contradiction (falsum) with the real world. This unfortunately is not possible to model. These restrictions belong to  $t_3$ . On the contrary, since it is up to the observer to either accept the excuse or reject the excuse, it cannot be tested in  $t_3$  if  $p$  actually caused the change in beliefs of the observer. This has to be tested in  $t_4$ . However, the two restrictions for  $p$  found in psychology belong to  $t_3$ . As stated before it is not possible to model these restrictions. Only for the restriction found in philosophy is it possible to propose

a model:  $[a \text{ xstit}]\pi$ .  $\pi$  denotes a situation in which agent  $a$  told  $p$ . Taking these two models together I propose the following definition: Agent  $a$  sees to it by excusing that  $\pi$  is stated:

$$[a \text{ exc}]\pi \cong [a \text{ xint}]\varepsilon \wedge [a \text{ xstit}]\pi$$

After this is done it should be checked if no contradiction occurs. I keep to words for this restriction. However, if no contradiction comes up it is possible the observer will change its view of the world (t4). But, if the excuse is accepted as such, it is up to the observer himself. This is not an objective decision and therefore it is hard or even impossible to model, as stated in the introduction of this thesis. It depends on the context, the relation between observer and excuser, and maybe even the emotional state of the observer. The only thing that can be stated is that if the excuse of agent  $a$  is accepted by the observer the view of the world of the observer is changed, which means that agent  $a$  is no longer held responsible for the bad conduct (t5). Therefore, I propose that t4 and t5 should be put together into one definition that says that observer  $o$  does not hold agent  $a$  responsible for seeing to it that  $\varphi$ . For this definition I also stick to human language.

The only problem left is that if the observer does not hold Agent  $a$  responsible for some other reason than the excuse made at t3, this is also true but the excuse made at t3 is not a proper excuse. Therefore an extra restriction is needed. An excuse is only a proper excuse if agent  $a$  is no longer held responsible only because he excused for his conduct. Again this is a counterfactual.

## Chapter 6

# Conclusions and Discussions

In the previous chapter all the different insights from psychology, philosophy and logic were integrated. That chapter was also concluding in a certain way. Therefore in the present chapter I will only give a summary of the overarching definition of excuses and highlight those parts that are of interest for discussion and further research.

It was noticed that psychology and philosophy both have a different focus on the concept of excuse. Philosophy is much more focused on conduct where psychology is focused on people: the excuser and the observer. Besides this a difference in perspective is that philosophy limits the concept of excuse much more than psychology does. To be able to compare the restrictions of both disciplines I had to narrow down the concept of excuses in psychology. In logic there is not much literature about the concept of excuses as a whole. However a lot of the restrictions found in psychology and philosophy were described in logic on their own. Due to the use of the restrictions found in other disciplines, biases of these disciplines were a threat to the chapter on logic. I tried to minimize this. However, I think that it is not possible to state that this thesis probed the concept of excuses in three different disciplines. It is better to state that it is probed in two disciplines and that logic is used to investigate the restrictions found in the other two disciplines. Therefore the role of logic is as important as psychology and philosophy since it can test hypotheses in a certain way. Due to the integration with logic a lot of questions have come up, which I will discuss while giving a summary of the found restrictions.

Both in psychology and philosophy it was noticed that there is a chronology in making an excuse. Some restrictions have to hold before others can be tested. The timetable exists of five steps. Some steps be met only one restriction, others hold more restrictions.

t1 holds all restrictions which have to do with the action done by agent *a*: the conduct has to be bad, it must be possible to commit the conduct intentionally and one of the following restrictions have to hold: coercion, mistake and unknowingly done. As discussed in the chapter on philosophy and in the previous chapter counterfactuals play an important role in these restrictions. So far, counterfactuals have not been included in XSTIT logic. Since XSTIT logic has much to do with knowledge, and counterfactuals are a major topic in epistemic philosophy, it should be interesting to do more research on how to incorporate counterfactuals in XSTIT logic.

t2 holds all restrictions about whether the observer holds agent *a* responsible: the observer belongs in some way to a different class of people than agent *a* and the observer knows agent *a* committed bad conduct at t1. For both restrictions has not been possible to make use of STIT logic. The first can be explained in set theory. For the second it might be possible to use STIT logic if an operator is designed which shows what people know about the past instead of the future. I think this is interesting to investigate.

t3 holds all restrictions on how agent *a* should excuse in order to make a proper excuse:

the excuser has to have the intention to excuse, agent  $a$  has to excuse by stating a proposition which will lead to the acceptance of the excuse on a later moment, and the excuse must not lead to a contradiction with the real world. For the first two restrictions I made a proposal on how to model it in XSTIT logic. The last restriction however cannot be modeled in XSTIT logic. The visualized world by agent  $a$  should be compared to the real world and the restriction holds only if no falsum occurs due to the part of the world, which has to do with the bad conduct.  $t_3$  is the actual moment in which is excused.

$t_4$  holds all restrictions for whether the view of the world of the observer is changed: the view of the world only changes if the observer accepts the excuse. It was not possible in this thesis to make restrictions on this since this is very subjective and depends on too many different aspects. Therefore  $t_4$  and  $t_5$  are discussed together.  $t_5$ , which holds the restrictions for whether agent  $a$  is no longer liable for the action, is only reached if the excuse is accepted as such by the observer. This acceptance has to be caused by the proposition stated by agent  $a$  in  $t_3$  otherwise the excuse does not count as a proper excuse. It should be interesting to probe further in psychology and philosophy for what reasons an excuse is accepted or not. After which it might be possible to probe how these restrictions could be incorporated in XSTIT logic.

Thanks to my research in the different disciplines I was able to create a semi formal model of all restrictions on excuses as stated in psychology, philosophy and or logics. I used all the formula's as I proposed them in this thesis. The model shows to which part of the timetable each proposed definition belongs. To know if an excuse is a proper excuse all restrictions have to be checked in order and start with the restriction about bad conduct. If a definition holds follow the arrow with yes, otherwise follow the arrow with no. See figure 6.1

By writing this thesis I found more questions than answers. I think it would be interesting to further investigate the concept on excuses in logic. Until now I only stated semi formal proves and noticed XSTIT is not developed far enough to describe the concept of excuses in total. I think using a certain concept of human behavior is a proper way to find out how complete a certain logic is. By finding restrictions on how to make a proper excuse it became clear that on some parts XSTIT logic was more specific on restrictions than human language, but on the other hand it also became clear that on some parts it was not possible to describe in XSTIT logic what we can describe in human language. Further research on the question formulated above will contribute to the further development of XSTIT logic.

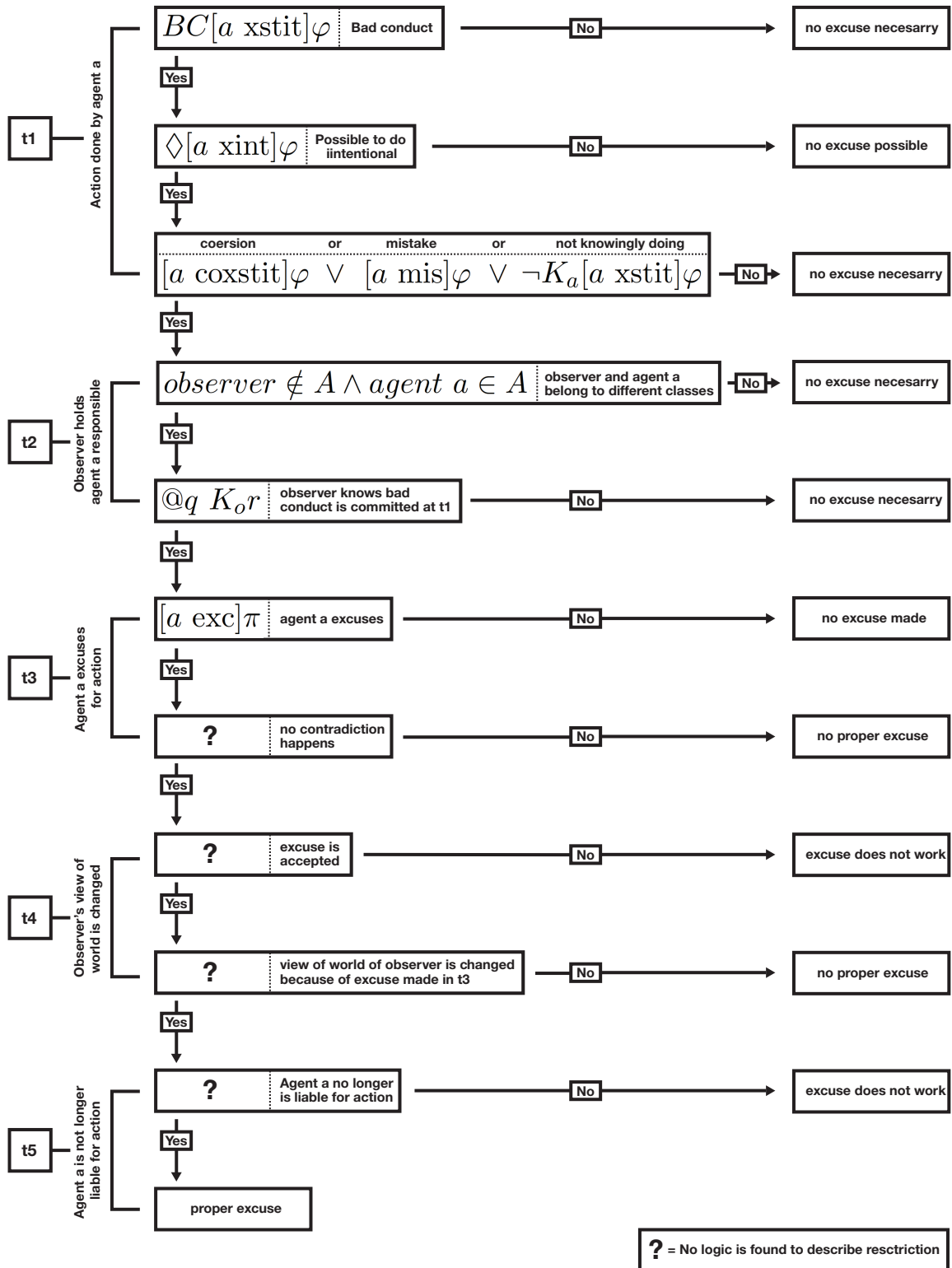


Figure 6.1: Model of excuses1.

# Bibliography

- [1] J. Austin, *A Plea for Excuses*, Proceedings of the Aristotelian Society 7 (1956) 1-30.
- [2] B. Baker, *The Excuse of Accident*, Ethics 93 (4) (1983) 695-708.
- [3] A. Baltag and S. Smets. *Dynamic belief revision over multi-agent plausibility models* In Proceedings LOFT 2006, 2006.
- [4] G. Boella, J. Broersen, L. van der Torre and S. Villata, *Representing Excuses in Social Dependence Networks*, AI\*IA 2009: Emergent Perspectives in Artificial Intelligence 5883 (2009) 365-374.
- [5] R. Brandt, *A Utilitarian Theory of Excuses*, The Philosophical Review 78 (3) (1969) 337-361.
- [6] M. Bratman, *Two Faces of Intention*, Philosophical Review 93 (1984) 375-405.
- [7] J. Broersen, *A Logical Analysis of the Interaction between ‘Obligation-to-do’ and ‘Knowingly Doing’*, in: L. v.d. Torre and R. v.d. Meyden (Eds.), Proceedings of 9th International Workshop on Deontic Logic in Computerscience (DEON’08), in: Lecture Notes in Computer Science, 5076, Springer, 2008, 140-154.
- [8] J. Broersen, *Deontic Epistemic Stit Logic Distinguishing Modes of Men Rea*, Journal of Applied Logic 9 (2011) 137-152.
- [9] J. Broersen, *Making a Start with the STIT Logic Analysis of Intentional Action*, Journal of Philosophical Logic 40 (2011) 399-420.
- [10] J. Broersen, *Modeling Attempt and Action Failure in Probabilistic Stit Logic*, in: T. Walsh (Ed.), Proceedings of Twenty-Second International Joint Conference on Artificial Intelligence (IJCAI 2011), 792-797.
- [11] H. Frankfurt, *Alternate Possibilities and Moral Responsibility*, The Journal of Philosophy 66 (23) (1969) 829-839.
- [12] J. Horty and N. Belnap, *The deliberative stit: a study of action, omission, and obligation*, Journal of Philosophical Logic 24 (6) (1995) 583-644.
- [13] J. van Lith - van Dis, *Hoorcollege 4-oktober*, in: Cursus kennis en wetenschap I, Universiteit Utrecht, Utrecht, 2012, pp. 5-7.
- [14] J. Montmarquet, *Culpable Ignorance and Excuses*, Philosophical Studies 80 (1995) 41-49.
- [15] B. Schlenker, T. Britt, J. Pennington, R. Murphy and K. Doherty, *The Triangle Model of Responsibility*, Psychological Review 101 (4) (1994) 632-652.



- [16] B. Schlenker, B. Pontari and A. Christopher, *Excuses and Character: Personal and Social Implications of Excuses*, *Personality and Social Psychology Review*, 5 (1) (2001) 15-32.
- [17] J. Searle, *Austin on Locutionary and Illocutionary Acts*, *The Philosophical Review* 77 (4) (1986) 405-424.
- [18] J. Searle, *The Structure of Illocutionary Acts*, in: A. Martinich (Ed.) *The Philosophy of Language*, Oxford University Press, New York, 2010, pp.146-156.
- [19] J. Searle, *The taxonomy of Illocutionary Acts*, in: A. Martinich (Ed.) *The Philosophy of Language*, Oxford University Press, New York, 2010, pp.157-170.
- [20] K. Sheldon and T. Schachtman, *Obligations, Internalization, and Excuse Making: Integrating the Triangle Model and Self-Determination Theory*. *Journal of Personality*, 75 (2) (2007).
- [21] R. Stalnaker, *Assertion*, in: A. Martinich (Ed.) *The Philosophy of Language*, Oxford University Press, New York, 2010, pp.196-208.
- [22] C. Snyder, R. Higgins and R. Stucky. *Excuses: Masquerades in Search of Grace*. New York, Wiley, 1983.
- [23] C. Snyder and R. Higgins, *Excuses: Their Effective Role in the Negotiation of Reality*. *Psychological Bulletin* 104 (1) (1988) 23-35.
- [24] P. Westen, *An Attitudinal Theory of Excuse*, *Law and Philosophy* 25 (2006) 289-375.