MODELING COUNTERFACTUAL STATEMENTS USING BAYESIAN NETWORKS: A CASE STUDY IN TURKISH

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS OF
THE MIDDLE EAST TECHNICAL UNIVERSITY
BY

MEHMET İLTERİŞ BOZKURT

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF COGNITIVE SCIENCE

APRIL 2023

Approval of the thesis:

MODELING COUNTERFACTUAL STATEMENTS USING BAYESIAN NETWORKS: A CASE STUDY IN TURKISH

Submitted by **MEHMET İLTERİŞ BOZKURT** in partial fulfillment of the requirements for the degree of **Master of Science in Cognitive Science Department, Middle East Technical University** by,

Prof. Dr. Banu GÜNEL KILIÇ
Dean, **Graduate School of Informatics** ____________________

Dr. Ceyhan TEMÜRCÜ
Head of Department, **Cognitive Science** ____________________

Assoc. Prof. Dr. Barbaros YET
Supervisor, **Cognitive Science Dept., METU** ____________________

Prof. Dr. Cem BOZŞAHİN
Co-Supervisor, **Cognitive Science Dept., METU** ____________________

Examining Committee Members:

Prof. Dr. Deniz ZEYREK BOZŞAHİN
Cognitive Science Dept., METU ____________________

Assoc. Prof. Dr. Barbaros YET
Cognitive Science Dept., METU ____________________

Prof. Dr. Özgür AYDIN
Linguistics Dept., Ankara University ____________________

Date: 24.04.2023

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : **Mehmet İlteriş BOZKURT**

Signature : ______________________

ABSTRACT

MODELING COUNTERFACTUAL STATEMENTS USING BAYESIAN NETWORKS: A CASE STUDY IN TURKISH

BOZKURT, Mehmet İlteriş

MSc, Department of Cognitive Science

Supervisor: Assoc. Prof. Dr. Barbaros YET

Co-Supervisor: Prof. Dr. Cem BOZŞAHİN

April 2023, 82 pages

Counterfactual statements are defined as conditional statements whose antecedents are false or unrealized (e.g., Anderson, 1951; Barwise, 1986; Ruhi et al., 2000; Pearl et al., 2016 among many others), which is one of the claims we challenge. The utilization and comprehension of counterfactual statements involve intricate cognitive mechanisms, including the ability to disregard actual events, the maintenance of both factual and counterfactual representations within mind, and the capacity to alternate between these representations. Counterfactuals have been extensively studied from a linguistic perspective, and more recently, they have been an important part of causality and causal modeling research with Judea Pearl at helm. The goal of this thesis is to contribute to this literature by combining these two perspectives in analyzing counterfactuals in Turkish, which we believe is a fresh approach that will shed new light on their nature. The research topics and questions include the following: whether the antecedent needs to be false or unrealized for one to be able to use a counterfactual construction, the use of *-sAydI*, a complex suffix in Turkish, and its relation to counterfactual interpretation, the difference between *-DIysA* and *-sAydI*, the rungs of Pearl & Mackenzie's (2018) *Ladder of Causation* that sentences involving the use of *-DIysA* and *-sAydI* fall, and the role of pragmatics or context in the interpretation of a counterfactual statement. The contributions of the thesis include being the first to have a clear focus on Turkish counterfactuals and making use of Bayesian networks to graphically represent counterfactual scenarios in Turkish.

Keywords: conditionals, counterfactuals, linguistics, causality, causal modeling

ÖZ

KARŞI OLGUSAL İFADELERİN BAYES AĞLARI KULLANILARAK MODELLENMESİ: TÜRKÇE ÜZERİNE BİR DURUM ANALİZİ

BOZKURT, Mehmet İlteriş

Yüksek Lisans, Bilişsel Bilimler Bölümü

Tez Yöneticisi: Doç. Dr. Barbaros YET

Ortak Tez Yöneticisi: Prof. Dr. Cem BOZŞAHİN

Nisan 2023, 82 sayfa

Karşı olgusal ifadeler, öncülleri yanlış veya gerçekleşmemiş olan koşullu ifadeler olarak tanımlanır (örn., Anderson, 1951; Barwise, 1986; Ruhi vd., 2000; Pearl vd., 2016 ve daha bir çoğu) ki bu, karşı çıktığımız iddialardan biridir. Karşı olgusal ifadelerin kullanımı ve kavranması, gerçek olayları göz ardı etme yeteneği, hem olgusal hem de karşı olgusal temsillerin zihinde tutulması ve bu temsiller arasında geçiş yapma kapasitesi dahil olmak üzere karmaşık bilişsel mekanizmaları içerir. Karşı olgusal ifadeler, dilbilimsel bir bakış açısıyla kapsamlı bir şekilde incelenmiştir ve son zamanlarda, Judea Pearl'ün öncülüğünde nedensellik ve nedensel modelleme araştırmalarının önemli bir parçası olmuştur. Bu tezin amacı, bu iki bakış açısını Türkçedeki karşı olgusal ifadeleri analiz etme konusunda birleştirerek bu literatüre katkıda bulunmaktır; bunun, karşı olgusal ifadelerin doğasına yeni bir ışık tutacak taze bir yaklaşım olduğuna inanıyoruz. Araştırma konuları ve soruları arasında, karşı olgusal bir yapının kullanılabilmesi için öncülün yanlış veya gerçekleşmemiş olmasının gerekip gerekmediği, Türkçede karmaşık bir ek olan *-sAydI* ekinin kullanımı ve karşı olgusal yorumla ilişkisi, *-DIysA* ve *-sAydI* arasındaki fark, *-DIysA* ve *-sAydI* kullanımını içeren cümlelerin Pearl ve Mackenzie'nin (2018) *Nedensellik Merdiveni*'nde bulundukları basamak ve bir karşı olgusal ifadenin yorumlanmasında edimbilimin veya bağlamın rolü yer almaktadır. Tezin katkıları arasında, Türkçe karşı olgusal ifadelere net bir şekilde odaklanan ve Türkçe karşı olgusal senaryoları grafiksel olarak temsil etmek için Bayes Ağlarından faydalanan ilk çalışma olmak sayılabilir.

Anahtar Sözcükler: koşullu ifadeler, karşı olgusal ifadeler, dilbilim, nedensellik, nedensel modelleme

Dedicated to

my dear mom Hafize
my sweet sister Tuğçe
my beloved spouse Zeynep

In Loving Memory of

my amazing grandfather Mehmet
my adorable grandmother Güler

ACKNOWLEDGMENTS

The successful completion of this thesis owes a great deal to the invaluable contributions of numerous individuals. While it is impossible to compile an exhaustive list and expressing feelings through writing is definitely not my forte, I would like to express my heartfelt gratitude to some of them.

I would like to start by conveying my deepest gratitude to my supervisor, Assoc. Prof. Dr. Barbaros Yet, and my co-supervisor, Prof. Dr. Cem Bozşahin, for their exceptional guidance, limitless patience, and steady support. Without our consistent three-way Zoom meetings during the early phases of molding this *-DIysA* and *-sAydI* idea Cem Hoca came up with into a full-on thesis and the invaluable face-to-face interactions with Barbaros Hoca in the later stages, this endeavor would not have possibly been accomplished. I truly consider myself lucky to have been their *supervisee*.

In addition to my supervisors, I would like to express my sincere appreciation to my committee members, Prof. Dr. Deniz Zeyrek Bozşahin and Prof. Dr. Özgür Aydın, for generously granting me a much-needed extension period. Their constructive critiques, enlightening recommendations, and priceless remarks were of immense assistance.

I would also like to thank Asst. Prof. Dr. Umut Özge and Dr. Ceyhan Temürcü for their understanding, particularly in relation to my TA responsibilities, during this entire process, and their encouragement in general.

I appreciate M. Hakan Güler and Sibel Gülnar for their continuous support in assisting students, including myself, throughout my time at the Informatics Institute and during the thesis submission process.

Thank you to my academia, *halı saha*, and gaming friends, and well, to all my friends in general. I appreciate each and every one of you, even if I choose to not mention you by name. The ones who played a significant role in keeping me motivated, you already know how much you mean to me.

I guess everyone longs for more moments with their departed loved ones. *Dede*, *Anneanne*, I yearn for your presence, that really is it, I simply miss you deeply. Coping with losing you both has been tough but I know you would want me to hang in there; so, I will continue forward, cherishing your memory every step of the way. Thank you.

Mom, I picture you trying to read and make sense of this paragraph putting all your Duolingo practice to work, and you look so cute! I extend my profound gratitude to you, who, despite having faced the profound loss of both your parents, enduring hardships of your own, and navigating tough life transitions, have been an unwavering pillar of support throughout my journey. Your enduring love, resilience, and belief in my abilities have been instrumental in my success. I am truly grateful for your steadfast presence and guidance, as I could not have achieved this milestone without your indomitable spirit by my side.

Sister, I am thankful for your immeasurable love. You have literally been there for me since the very beginning, that is to say, the date of my birth, so, I cannot overstate the significance of your support. You had a hand in all my major decisions; in this case, you embraced this endeavor as if it were your own, sharing in both the highs and lows. Notwithstanding the fact that you had your hands full, you always had time to motivate me, help me plan, and ask how I was doing in general. I am forever touched by your faith in me, and I know that this thesis would not have been completed if you had not shared your experience with me. I know we have had our differences at times but we always found a way to reconcile; you were, you are, and I hope you will be there for me. Also, I cannot overlook the contributions of Serkan, your husband and my brother-in-law; he provided direct support to me in every possible way and indirect support through always being there for you. I am happy to have such an incredible sister and I feel very fortunate to have met you, Serkan!

Zeynep, if someone had told me that we would be united in matrimony by the time I handed in this thesis, I would not have believed them. How could I have fathomed that the timid girl I nervously sat next to on our first day of high school would become my beloved wife? Yet, here we are, we will have embarked on this new chapter of our lives for only a day or two when I submit this work. We have shared a remarkable journey of ~10 years together, overcoming the challenges of a long-distance relationship and persevering through it all. I am forever grateful for your belief in me and your understanding during the countless hours I spent acting like hanging on in quiet desperation is the *Turkish* way. It was your patience and encouragement that sustained me during moments of doubt and exhaustion and provided me with a sense of security and comfort. I eagerly anticipate the years we will spend together, building a future filled with love and shared accomplishments. Thank you for being the anchor that grounds me and keeps me away from drifting into nothingness, and the constant source of inspiration in my life. I am truly blessed to have you by my side.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

1/2/3SG	First/second/third person singular
1/2/3PL	First/second/third person plural
ABL	Ablative
AOR/NEGAOR	Aorist/negated aorist
AUX	Auxiliary verb
BN	Bayesian network
CAUS	Causative
CPT	Conditional Probability Table
COGS	Cognitive Science
COND/CONDCOP	Verbal/copular conditional
DAG	Directed Acyclic Graph
EVID/EVIDCOP	Verbal/copular evidentiality
FUT	Future
GM	Generalizing modality
IMP	Imperative
IMPF	Imperfective
INT	Interrogative
OBLG	Obligative
PASS	Passive
PAST/PASTCOP	Verbal/copular past
PF	Perfective
PROG	Progressive
PSB	Possibility
METU	Middle East Technical University
NEG	Negation
THEY/THEIR/THEM	Pronouns used in Turkish to English translations throughout the thesis instead of *s/he* / *her/his* / *her/him*

CHAPTER 1

1. INTRODUCTION

“Time. Space. Reality. It’s more than a linear path. It’s a prism of endless possibility, where a single choice can branch out into infinite realities, creating alternate worlds from the ones you know. I am The Watcher. I am your guide through these vast new realities. Follow me, and ponder the question... What if?”

The Watcher - Disney+ animated series “What If...?” (2021)

Who among us has not pondered the potential consequences of an event unfolding differently from how it actually did? The said event can be about your personal life as in (1), the circumstances of a football match as in (2), or the history of a country as in (3). Contemplating different possibilities that could have occurred instead of what actually happened is a fundamental part of human thinking and emotion (Epstude & Roese, 2008). These types of thoughts are known as counterfactual thoughts.

(1) [If I had eaten a healthier breakfast], I would have had more energy throughout the day.
[Daha sağlıklı bir kahvaltı yapsaydım] gün içinde daha fazla enerjim olurdu.

(2) [If there had been VAR in 2013], Galatasaray would have knocked Real Madrid out.
[2013’te VAR olsaydı] Galatasaray Real Madrid’i elerdi.

(3) [If Atatürk had lived to see his 90s], Turkey would have been a very different place.
[Atatürk 90’lı yaşlarını görebilseydi] Türkiye çok farklı bir yer olurdu.

In all three examples above, the first part of the sentence is simply not the case, or to be more precise, the events have transpired differently: in (1), the speaker did not eat a healthy breakfast, and they were not energetic on the given day; in (2), VAR (Video Assistant Referee) was first used in full in 2018 World Cup, and Galatasaray was eliminated from the Champions League in April 2013; and in (3), Atatürk sadly passed away in 1938 at the age of 57, so he did not get to see his 90s.

Counterfactual statements are defined as conditional statements in which the antecedent is false or unrealized (Anderson, 1951; Barwise, 1986; Ruhi et al., 2000; Pearl et al., 2016), which is a claim we will argue against in later chapters. The utilization and comprehension of counterfactual statements necessitate the engagement of intricate cognitive mechanisms. Specifically, the processing of counterfactuals necessitates the execution of three cognitive operations: (i) the ability to disregard actual events, (ii) the maintenance of both factual and counterfactual representations within cognitive architecture, and (iii) the capacity to alternate between these representations (Beck et al., 2009; Drayton, Turley-Ames, & Guajardo, 2011; Guajardo, Parker, & Turley-Ames, 2009 as cited in Yarbay Duman, 2015).

Counterfactuals have long been a topic of interest in many fields:

Law: They have been used in legal decision-making to evaluate liability and consider the potential consequences of different legal outcomes. One notable example of this is the *but-for* test, which is used to determine causation in tort law. According to this test, the defendant's actions must have been the *but-for cause* of the plaintiff's injury or damage for the defendant to be held liable (Feinberg, 1970). This type of counterfactual analysis is essential for determining legal responsibility and ensuring that individuals are held accountable for the harm they cause.

History: They have long been a central tool in the study of history, as they allow historians to consider the potential impacts of different events and decisions on the course of history. For instance, counterfactual analysis has been used to explore what might have happened if major historical figures had not died prematurely or if key events had not occurred. This type of analysis can provide insight into the complex web of causes and consequences that shape historical events, and help historians to better understand the forces that have shaped the world in which we live. One notable example of this type of analysis is Ferguson (1997), which explores a range of counterfactual scenarios and considers their potential impacts on history.

Economics/Econometrics: They have been used to evaluate the alternative outcomes of different economic policies or decisions. For example, economists have employed counterfactual analysis to consider how the economy might have fared if the Federal Reserve had not lowered interest rates during the 2008 financial crisis (Bernanke et al., 1999), or if the United States had not implemented a minimum wage (Card & Krueger, 1995). Counterfactual analysis is also used in the field of econometrics, where it is used to identify the causal relationship between different economic variables (Wooldridge, 2010).

Philosophy: They have been widely studied as a way of understanding counterfactual reasoning and its role in various fields, such as decision-making and moral responsibility.

For example, Lewis (1973a) presents a formal system for analyzing counterfactual statements and offers a comprehensive theory of counterfactual dependence, and Plantinga (1974) examines the role of counterfactuals in modal metaphysics and the concept of possible worlds. Other notable contributions to the study of counterfactual reasoning include Marcus (1961), which explores the logical properties of counterfactual statements, and Lewis (1973b), which investigates the relationship between counterfactuals and comparative possibility.

Psychology: They have been used to study how people think about and evaluate events and their outcomes. For example, research has shown that people who engage in more counterfactual thinking after a negative event tend to feel more regret and less acceptance, while those who engage in less counterfactual thinking tend to feel more acceptance and less regret (Roese & Olson, 1995). In addition to its role in coping and decision-making, counterfactual thinking has also been studied in relation to creativity, problem-solving, and social comparison (Kahneman & Tversky, 1982).

Sociology: They have been used to explore the potential consequences of different social policies or interventions. For example, sociologists have used counterfactuals to examine the potential outcomes of different educational policies on student achievement (Hanushek, 2011), the impact of alternative policies on crime rates (Ehrlich, 1996), or what might have happened if different social programs had been put in place to address poverty or inequality (Manski, 2008).

Artificial Intelligence: They have been used to help machine learning models better understand and predict the outcomes of different actions. For example, machine learning models have been developed that use counterfactuals to consider what would have happened if they had made a different decision in order to improve their future decision-making (Swaminathan & Joachims, 2015). Another example of the use of counterfactuals in artificial intelligence is in the development of counterfactual fairness algorithms, which aim to reduce bias in machine learning models by considering what the model's predictions would have been for different groups of people if certain characteristics (such as race or gender) had been different (Kusner et al., 2017).

Linguistics: They have been studied as a way of understanding how language is used to convey counterfactual meaning. For example, research has shown that different languages have unique strategies for expressing counterfactuals, such as using specific verb forms or modal auxiliaries (Iatridou, 2000; von Fintel, 1994). Linguists have also studied the role of counterfactuals in the interpretation of conditional statements, as well as the relationship between counterfactuals and tense and aspect (Kratzer, 1981).

Causality/Causal modeling: They have been used to understand the relationships between different variables and the outcomes that they produce. For example, researchers may use counterfactual analysis to consider what might have happened if a certain variable had been changed in order to understand the causal relationship between that variable and the outcome in question (Pearl, 2009).

Counterfactuals have already been extensively studied from a linguistics perspective (e.g., Anderson, 1951; Byrne, 2016; Iatridou, 2000; Yarbay Duman et al., 2015, 2016, among others), and more recently, they have been an important part of causality and causal modeling research with Judea Pearl at the helm (e.g., Pearl, 2009; Pearl et al., 2016; Pearl & Mackenzie, 2018, among others). Taking into account the interdisciplinary nature of cognitive science, the goal of this Cognitive Science MSc thesis is to contribute to this literature by combining these two perspectives in analyzing counterfactuals in Turkish, which we believe is a fresh approach that will shed new light on their nature.

1.1. Research Questions

The present study aims to answer the following questions:

1. Does the antecedent need to be false or unrealized for one to be able to use a counterfactual construction?
2. Does the use of *-sAydI*, a complex suffix that is a combination of the conditional marking verbal suffix *-sA* and past tense marking copula *-(y)dI*, in a Turkish sentence always lead to a counterfactual interpretation?
3. What happens when *-DI* precedes *-(y)sA* and what happens when it is vice versa? In other words, what exactly is the difference between *-DIysA* and *-sAydI*?
4. On which rung of Pearl's Ladder of Causation do the sentences which involve the use *-DIysA* and *-sAydI* fall?
5. Does pragmatics, or to be more precise, context play a role in the interpretation of a counterfactual statement?

1.2. Contributions of the Thesis

The contributions of the present study are as follows:

1. The main focus of the previous research that touched upon Turkish counterfactuals was actually conditional statements in general (e.g., Can Bakırlı, 2010; Göksel & Kerslake, 2005), and counterfactuals were only briefly mentioned; this work is the first to have a clear focus on them.
2. Influenced by the causal modeling research, this is the first work to make use of Bayesian networks (BNs) to graphically represent counterfactual scenarios in Turkish.

1.3. Organization of the Thesis

This thesis comprises five chapters, with this subchapter being the last one in the introductory one. Here, we introduced counterfactuals and touched upon the research fields in which they are used, and presented the research questions and the contributions of the thesis.

Chapter 2 is on conditionals and counterfactuals in Turkish. Here, the literature on conditionals and counterfactuals in general and Turkish conditionals and counterfactuals are reviewed. This chapter introduces the reader to the aforementioned linguistics part of our approach.

Chapter 3 is on the graphical representation of counterfactuals. Here, the focus is on causal inference, and BNs and the way they are used to represent counterfactuality are presented. This chapter introduces the reader to the aforementioned causality part of our approach.

Chapter 4 describes our approach that combines linguistics and causal modeling. Here, example use cases of the said approach that focus on the importance of context in interpreting counterfactuals are presented.

Chapter 5 comprises two parts: a conclusion that discusses and recapitulates the main points of the present study and future work.

CHAPTER 2

2. ON CONDITIONALS, COUNTERFACTUALS, AND TURKISH

"It doesn't matter. [takes a swig of Macallan] Even if the pictures are fake. I can't take it anymore. [turns away] Do you have any idea what a curse it is to have perfect 20/20 hindsight? [takes another swig] As soon as something bad happens, I immediately know how it could have been avoided. I can't take it anymore!"

Captain Hindsight - South Park S14E12: Mysterion Rises (2010)

This chapter focuses on existing research on conditionals and counterfactuals in general and Turkish conditionals and counterfactuals, and it provides an overview of the relevant linguistic concepts. The chapter introduces the reader to the *linguistics* part of the approach that is to be introduced in Chapter 4.

2.1. An Overview

Conditional structures express a conditional relationship between two events, often referred to as the *antecedent* and the *consequent*, or the *protasis* and the *apodosis*. According to Greenberg's (1963) Word Order Universal 14, the default order for a conditional is antecedent+consequent, or protasis+apodosis, but many languages allow the antecedent to be placed at the beginning or end of the sentence (Comrie, 1986), and the function of a conditional may vary depending on the order of the antecedent and the consequent. In English, conditionals are signaled by the use of an *if* in the antecedent and an optional *then* in the consequent, as in (4) (*then* is pretty interesting per se but we will not provide a lengthy discussion about it, see Iatridou, 1993, instead).

(4) If he studies hard, (then) he will pass the exam.
(or *he will pass the exam if he studies hard*)

The antecedent is the event that must occur in order for the consequent to take place. The antecedent and the consequent are evaluated based on the proposition presented in the former, which determines the conditions under which the proposition expressed in the latter is true (Bhatt & Pancheva, 2017) for example, in (4), the if clause, *if he studies hard*, expresses the condition of a certain individual's studying hard, while the main clause, *he will pass the exam*, expresses the expected consequence of that effort, namely, the said individual's passing of the exam.

Conditionality is a concept that is studied in various disciplines (Oktar & Can, 2016) and conditionals are believed to be essential for understanding language because they are multifunctional (Athanasiadou, 1997). They have been classified as a subtype of adverbial clauses that express time, reason, and concession (Thompson et al., 2007). Most studies on these constructions, which are encoded by specific markers in various languages, only examine them at the sentence level rather than within a contextual framework, as we believe they should.

There are other ways to convey conditionality in English in addition to the canonical form of *if X, (then) Y,* as in (5-11) (Cullicover & Jackendoff, 1999; Iatridou, 2013; von Fintel, 2011; Stump, 1985).

(5) Had I known that you were sick, I would have visited you.
(=*If I had known you were sick, then I would have visited you.*)

(6) Ignore your homework and you will fail.
(=*If you ignore your homework, you will fail.)*

(7) Standing on a chair, he can reach the ceiling.
(=*If he stands on a chair, he can reach the ceiling.*)

(8) He was pushed or he would not have fallen down the cliff.
(=*If he was not pushed, he would not have fallen down the cliff.*)

(9) Without you, I would be lost.
(=*If you did not exist/were not in my life, I would be/would have been lost.*)

(10) The more he waits, the angrier he gets.
(=*If he continues to wait/waits longer, he gets angrier.*)

(11) To have danced in the park would have been wonderful.
(=*If we had danced in the park, it would have been wonderful.*)

2.1.1. Types of Conditionals

There are four types of conditionals that have been identified in academic literature. The first of these are *speech act conditionals*[1], also known as *biscuit conditionals* or *relevance conditionals*, which are characterized by a lack of stated conditions under which the consequent is true (von Fintel, 2011). An example of this type of conditional can be seen in (12) (Austin, 1956).

(12) There are biscuits on the sideboard if you want them.

---

[1] *Speech acts* themselves are outside the scope of this work, see Austin (1975) and/or Searle (1969).

In this type of conditional, the antecedent does not specify the circumstances in which the proposition expressed in the consequent is true; instead, it establishes its relevance from the perspective of the speech act. These conditionals can be thought of as containing an implicit performative clause that embeds the main clause, with the performative serving as the true consequent (Bhatt & Pancheva, 2017). For example, in (12), if you are hungry, then it is relevant for you to know that there are biscuits on the sideboard, the antecedent establishes the relevance of the proposition in the consequent (that there are biscuits on the sideboard) for the listener.

Speech act conditionals are often given less attention in discussions of conditionals (but of course, see DeRose & Grandy, 1999; Siegel, 2006; and Predelli, 2009 as cited in von Fintel, 2011, for notable exceptions).

The second type, which is also not in the limelight when it comes to the literature on conditionals, is exemplified by the likes of (13), in which the antecedent introduced by someone else is restated (von Fintel, 2011).

(13) If you're so clever, why don't you do this problem on your own?

They are called *factual conditionals* by Iatridou (1991) in her PhD dissertation on conditionals, who maintains that their antecedent is assumed to be true, and *premise conditionals* by Haegeman (2003), who suggests that while the antecedents clause in standard conditionals are integrated into the speech act of the matrix clause, the antecedent in premise conditionals have a separate illocutionary force.

The primary focus on conditionals is on the remaining types, typically referred to as *indicative conditionals* and *subjunctive/counterfactual conditionals* (Edgington, 1997; Dancygier, 1988; Givón, 1995; Wu, 1993 as cited in Declerck & Reed, 2001). The focus of the current thesis will be on these two, as well.

2.1.2. Indicative & Subjunctive/Counterfactual Conditionals

The difference between the two is often illustrated using the Oswald & Kennedy minimal pair from Adams (1970). Below, we adapt this US-based example to Turkey, specifically to the assassination of Abdi İpekçi, a newspaper editor, by Mehmet Ali Ağca on February 1st, 1969. (14) exemplifies indicative conditionals and (15) subjunctive/counterfactual conditionals.

(14) If Ağca did not kill İpekçi, someone else did.
(15) If Ağca had not killed İpekçi, someone else would have.

The main distinction between their meanings is that indicative conditionals convey that the truth of the antecedent is uncertain, while subjunctive conditionals convey that the antecedent is false. von Fintel (2011) claims that the majority of people would accept (14) as true, given their knowledge of İpekçi's assassination, while only a subset of people would accept (15).

The meaning of the indicative conditional aligns closely with the *Ramsey Test* (Ramsey, 1931), which posits that if two individuals are discussing *If p will q?* and both have doubts about *p*, then they are hypothetically adding *p* to their knowledge and debating about *q*. Given the knowledge of İpekçi's assassination and a small uncertainty about whether Ağca was the assassin, hypothetically adding *Ağca did not kill İpekçi* to our knowledge would support the idea that someone else killed İpekçi and thus, (14) is considered true. However, (15) is evaluated differently, as it invites us to consider how things would have turned out if Ağca had not killed İpekçi, and most people do not have a clear idea of what would have happened, leading to the rejection of (15) (von Fintel, 2011).

While the terms *indicative* and *subjunctive* suggest that the distinction is based on grammatical mood, and the term *counterfactual* implies that the second type involves a contrary-to-fact assumption, von Fintel (2011) puts forward the idea that neither terminology is entirely accurate. The outward difference between the two types of conditionals lies in their tense/mood/aspect syntax, but it is not consistently a difference in indicative versus subjunctive mood. In languages that have a past subjunctive, such as German, the antecedents of subjunctive conditionals appear in the subjunctive mood, but if a language does not have a past subjunctive, another form is used. English uses an indicative antecedent with an additional layer of past tense morphology, and that reliably signals subjunctive conditionals.

Several prominent theories have been proposed to explain the nature of conditionals and counterfactuals, but they are not in the scope of this work, which aims to use the graphical representation techniques of causal modeling research to analyze counterfactuals in Turkish. So, to get an overview of conditionals and counterfactuals from a cross-linguistic perspective, we invite the reader to consult Bhatt & Pancheva (2017) for their syntax and von Fintel (2011, 2012) for their semantics. The Stanford Encyclopedia of Philosophy entries on *Indicative Conditionals*, *The Logic of Conditionals*, and *Counterfactuals* are also immensely stimulating (compiled by Edgington, 2020; Egré & Rott, 2021; Starr, 2022, respectively).

2.2. Turkish Conditionals and Counterfactuals

This subchapter will focus on by far the most comprehensive account of conditionals in Turkish, Göksel & Kerslake (2005), a grammar book which provides a thorough and orderly explication of the linguistic structure at all levels and one that has been a seminal reference for almost any work done on Turkish. Another important work, Can Bakırlı (2010)[2], is a PhD dissertation on conditionals in Turkish, in which an eight-novel corpus of detective fiction was analyzed through the lens of Declerck & Reed's (2001) approach to conditional clauses, as well as the framework proposed by Chang-Bong (2001, as cited in Can Bakırlı, 2010) which is based on Prince (1992, as cited in Can Bakırlı, 2010).

[2] Note that the original work is in Turkish.

2.2.1. The Verbal Suffix -sA and the Copular Marker -(y)sA

A canonical conditional construction in Turkish involves the use of either the verbal suffix -*sA*, as in (16), or the copular marker -*(y)*sA[3], as in (17), in the antecedent (Göksel & Kerslake, 2005). These two markers were distinguished as: i) *conditional* and *suppositional* by Deny (1921), *counterfactual* and *factive conditional* by Lees (1962), and *subjunctive* and *indicative* by Kuruoğlu (1986) and Sezer (2001) (Deny, 1921; Lees, 1962; Kuruoğlu, 1986 as cited in Sezer 2001). Exhibiting an antecedent+consequent order by default (Can Bakırlı, 2010), Turkish is congruent with the aforementioned Word Order Universal 14 (Greenberg, 1963).

(16) [Ödevini yap-**sa**-n] iyi olur.
do-COND-2SG

It would be good [if you did your homework].

(17) [Ödevini yap-tı-**ysa**-n] korkacak bir şeyin yok.
do-PAST-CONDCOP-2SG

You have nothing to fear [if you have done your homework].

Göksel & Kerslake (2005) argue that the selection of a conditional marker is contingent upon the function of the conditional clause within the context of the sentence as a whole. Following that, they introduce three primary functional types of conditional sentences: (i) *predictive conditionals*, (ii) *knowable conditions*, and (iii) *universal conditionals*.

2.2.2. Predictive Conditionals

The predictive type of conditional sentence posits that if a particular event occurs (or has occurred), another event will follow (or would have followed) as a result. The first event is presented as the cause or enabling factor of the second event. Four potential markings exist for the subordinate clause of a predictive conditional: (a) -*(y)sA*, (b) -*sA*, (c) -*sAydI*, and (d) -*sAymIş*. (a) conveys a condition that is deemed capable of fulfillment, (c) denotes a condition known to be incapable of fulfillment, while (b) and (d) are equivocal in this regard (Göksel & Kerslake, 2005).

a) -(y)sA

-*(y)sA* is commonly affixed to the aorist -*(A/I)r* or its negated version, -*mAz*. The conditional clauses that utilize the aorist can be classified into two categories. The first category, which is the more apparent *predictive* form, is referred to as an *open conditional*. In this type, both the stated condition and its consequent are situated in the future. The

[3] -*(y)sA* has a free-standing alternative: *ise*, which is used scarcely in modern standard Turkish.

second category encompasses two closely related forms of statement, namely the *generic* and *habitual* forms. In this category, the predictability pertains not to the relationship between a pair of future events, but rather to the recurring pattern of events. When the verb in the main clause is also in the aorist form, there is a possibility for ambiguity between the two categories (Göksel & Kerslake, 2005).

In open conditionals, the speaker does not have knowledge of whether the proposed condition will be fulfilled but rather regards it as a realistic possibility rather than a hypothetical one. In instances where the main clause is a statement, the verb is marked by either the future marker *-(y)AcAK*, as in (18), or the aorist *-(A/I)r* or *-mAz*, as in (19). The use of the future marker indicates that the event is considered an inevitable outcome following the satisfaction of the condition, potentially due to a prior decision, whereas the use of the aorist in the main clause implies that the consequence of the condition being fulfilled is seen as an assumption or likelihood (Göksel & Kerslake, 2005).

(18)	[Projesini zamanında	teslim et**-mez-se**] dersten	kal**-acak**.
		turn.in/submit-NEGAOR-CONDCOP	fail-FUT

[If they do not turn their project in on time], they will fail the course.

(19)	[Final sınavına yeterince iyi	çalış**-ır-sa]** dersi kolayca	geç**-er**.
		study-AOR-CONDCOP	pass-AOR

[If they study well enough for the final exam], they will pass the course easily.

In contrast to other forms of predictive conditionals, where the main clause is restricted to being either a statement or a question, in open conditionals, the main clause can also take the form of a volitional utterance, as in (20), in which the speaker is talking about a newly-wed couple, Ali and Esra, whose wedding they failed to attend.

(20)	[Ali veya Esra'yı görürsen] onlara tebriklerimi	**ilet**.
		extend[IMP-2SG]

[If you see Ali or Esra], extend my congratulations to them.

If an open conditional construction appears within a narrative, and the situation being described was in the future at that point in the story, the past copula *-(y)DI*[4] is added to the main clause, as in (21).

(21) Veli gerçekten zor durumdaydı; [kendisine söyleneni yapmazsa] işini kaybedebilir**-di**.

[4] *-(y)DI*, like *-(y)sA*, has a free-standing alternative: *idi*, which is, like *ise*, used scarcely in modern standard Turkish.

Veli was in a really tough spot; he might lose his job [if he did not do what he was told].

Generic conditionals pertain to the inherent characteristics or actions of a class of entities, and are regularly articulated using the aorist form. In conditional sentences, there exists the potential for ambiguity between a generic reading and an open predictive interpretation. In (22a), *kedi* (cat) and *köpek* (dog) are understood in a generic sense, whereas in (22b) they are perceived as references to specific individuals of the species (Göksel & Kerslake, 2005).

(22) [Kediyle köpek bir arada büyüt**-ül-ür-ler-se**] dövüşmezler.
bring.up-PASS-AOR-3PL-CONDCOP

a. *[If a cat and a dog are brought up together], they do not fight.*
b. *[If the cat and the dog are brought up together], they will not fight.*

Habitual conditionals make generalizations about the behavior or qualities of specific entities or groups. Non-past habituals typically feature the progressive marker *-(I)yor* in the main clause, which fully resolves the ambiguity from open predictive conditionals, as in (23), and when a habitual conditional sentence contains a past time reference, the main clause is denoted by the past tense marking, while the conditional clause remains unmodified, as in (24) (Göksel & Kerslake, 2005).

(23) [Ali doğum gününü unutursa] Esra çok kız**-ıyor**.
get.angry-PROG

Esra gets very angry [if Ahmet forgets her birthday].

(24) [Vekil bir danışmanından memnun kalmazsa] onu sür**-dür-ür-dü**.
exile-CAUS-AOR-PASTCOP

[If the deputy was not pleased with one of his advisors], he would have him exiled.

-(y)sA also occurs in compound forms attached to the auxiliary verb *ol-* and the aorist. In open conditionals, i) *-mIş olursa* denotes that the event in the conditional clause is imagined as being completed before a specified reference point, as in (25), ii) *-(I)yor olursa* presents the event as being ongoing at a certain reference point, as in (26), and iii) *-(y)AcAk olursa* can indicate that an event is about to happen, as in (27), or that the chronological sequence of main clause event following conditional event is the same as with aorist + *-(y)sA*, as in (28). In habitual conditionals, the functions of these compound forms are clearly defined, with one exception: the function of *-(y)AcAk olursa* that is similar to aorist + *-(y)sA* does not occur (Göksel & Kerslake, 2005). (29), (30), and (31) all refer to a hypothetical situation that has not been experienced but will be encountered if experienced.

(25) [Anneleri gel-me**-miş ol-ur-sa**] çocuklarla kal**-abil-ir mi-siniz**?

come-NEG-PF AUX-AOR-CONDCOP stay-PSB-AOR INT-2SG/PL

[If their mother has not come back (by then)], would you be able to stay with the children?

(26) [Duruşma saat 3'te hâlâ devam ed**iyor olursa**] sinirli bir şekilde kalkıp giderim.

[If the trial is still going on at 3 PM], I will get up and leave in anger.

(27) [O güzel arabanı sat**acak olursa**n] bana bir alo de.

[If you are (ever) about to sell/decide to sell that beautiful car of yours], give me a call.

(28) [Veli sonradan fikrini değiştir**ecek olursa**] bunu hoş karşılamalıyız.

[If Veli should change his mind later], we must welcome it.

(29) [Yemeği Ali hazırla**mış olursa**] masayı Esra kuruyordu.

[If Ali had prepared/cooked the meal], Esra would set the table.

(30) [Hiç ara vermeden çalış**ıyor olursa**k] genelde yemeği dışarıdan söylüyoruz.

[If we are studying without any breaks], we usually order delivery.

(31) [Veli okuldan sonra spora gid**ecek olursa**] yanına uygun kıyafetler alıyor.

[If Veli is going to the gym after school], he takes appropriate clothes with him.

It is worth noting that certain forms of open or recurrent conditional clauses can be articulated without the utilization of the aorist tense. Specifically, when the conditional clause possesses a nominal predicate, the conditional copula may be appended directly to said predicate, without the need for the copular verb *ol-*; as such, according to Göksel & Kerslake, 2005), the following forms are interchangeable: *olursa* and *-(y)sA/varsa*, as well as *olmazsa* and *değilse/yoksa*. The examples Göksel & Kerslake (2005) use to display the said interchangeability in open conditionals without aorist marking are (32) and (33). Additionally, when the conditional clause comprises a verbal predicate marked with *-mIş* or *-(I)yor*, the conditional copula may be directly attached to the predicate without the formation of a compound verb form.

(32) [Yarın ev fazla soğuk **ol-maz-sa/değil-se**] burada çalışabiliriz.

be-NEGAOR-CONDCOP/not-CONDCOP

Tomorrow [if the house is not too cold] we can work here.

(33) [Başka işim **olmazsa/yoksa**] gelirim.

I will come [if I have no other commitments].

Since one of the main aims of this work is to highlight the importance of contextualization when evaluating conditionals/counterfactuals, we believe the aforementioned interchangeability is not entirely clear-cut. We believe the use of *değilse* in (32) and *yoksa* in (33) to be *marked*, or even *almost infelicitous*, without a context. To exemplify this claim: the only way for a native speaker of Turkish to use *değilse* in (32) is if they are just about to check a forecast to see how the weather will be the next day, as in (34), and similarly, the only way for *yoksa* to be used in (33) is if they will check their calendar to see if they will be otherwise occupied, as in (35). In other words, there needs to be a certain contextual enrichment that licenses the future references.

(34) [Yarın fazla soğuk değilse] burada çalışabiliriz, **hava durumuna bir bakalım**.
[If it is not too cold tomorrow] we can study here, ***let's check the weather forecast.***

(35) [Başka işim yoksa] gelirim, **takvimime bir bakayım**.
I will come [if I have no other commitments], ***let me take a look at my calendar.***

b) *-sA*

The utilization and the spectrum of significations of the verbal suffix -*sA* in conditional clauses exhibit a certain degree of intricacy. The determining factor in the interpretation of this suffix in any specific conditional sentence is the presence or absence of the past copula -*(y)DI* in the main clause. In instances where the main clause does not incorporate -*(y)DI*, the condition expressed in the subordinate clause is deemed to possess the potential for realization but is presented more as a hypothetical probability rather than one which the speaker perceives as being situated in the actual future. As a result of this, the verb in the main clause is nearly always in the aorist form (Göksel & Kerslake, 2005).

-*sA* is frequently employed when discussing actions that are being contemplated in a detached or abstract manner (Göksel & Kerslake, 2005). An example of this usage can be observed in sentence (36), which may serve as an initial contribution to a discourse concerning the means of arriving at a different city punctually for a specific meeting. -*sA* is also utilized in instances where the utterance calls into question the credibility or validity of something that has been stated or presupposed, or the suitability of an action that is planned or currently underway (Göksel & Kerslake, 2005). As in (37), this usage of the suffix serves to challenge the assumed state of affairs and can express a degree of skepticism or doubt.

(36) [Sabah 5'teki uçağa bin-**se**-k] 11'deki toplantıya yetiş-ebil-**ir** mi-yiz acaba?
take-COND-1PL make.it-PSB-AOR INT-1PL

Would we, I wonder, be able to make it to the 11 AM meeting (on time) [if we took the 5 AM flight?]

(37) [5'teki uçağa bin**se**k] yetişemez miyiz?

Would not we be able to make it (on time) [if we took the 5 AM flight]?

Another pervasive utilization of the *-sA* conditional construction is to convey a typically favorable assessment of a particular prospective action, frequently accompanied by a strong implication that the speaker desires or anticipates this action to be executed (Göksel & Kerslake, 2005). (38), which exemplifies this usage, serves to express encouragement for a certain cleaning to be done and/or indicates the speaker's expectation for the proposed action to be carried out.

(38)	[Bugün burada temizlik	yap-ıl-**sa**] iyi olur.	
		do-PASS-COND	

It would be good [if some cleaning were done here today].

In instances where the conditional clause is distinguished by the presence of the verbal suffix *-sA* and the main clause contains the past copula *-(y)DI*, the sentence conveys a counterfactual meaning; that is, it expresses a situation that contradicts the actual state of affairs. It is worthy of note that in this modal function as a marker of counterfactuality, *-(y)DI* does not exclusively pertain to past time, as in (39), in which the speaker does indeed have a lot to do, and they will not be joining the party the next day.

(39)	[Yapacak çok işim	ol-ma-**sa**] ben de yarınki partiye	gel-ir-**di**-m.
		be-NEG-COND	come-AOR-PASTCOP-1SG

I would come to the party tomorrow, too, [if I did not have a lot to do].

The utilization of *-sA* to express counterfactuality primarily occurs in conjunction with verbs that signify states, such as *ol-* (be or exist) and *bil-* (know), as in (40) (Göksel & Kerslake, 2005). With event verbs, counterfactuality is typically conveyed either by *-sAydI* or by compound forms with *olsa*; however, in instances where the main clause is of an evaluative nature, event verbs can also be employed in this pattern. In such utterances, the counterfactuality is less absolute than in similar sentences with *-sAydI* in that the possibility that the desired event may yet transpire in the future is not completely precluded.

(40) [O genç topçu dünkü maçta oyna**sa**] kesin fark yaratır**dı**.

[If that young baller (referring to a footballer) had played in yesterday's game], he would have definitely made a difference.

c) -sAydI [conditional marker -sA + past copula -(y)dI]

According to Göksel & Kerslake (2005), conditional clauses distinguished by the presence of *-sAydI* are consistently characterized by counterfactual meaning (which is a claim we will be arguing against in Chapter 4) and, as such, always feature the past copula in their main clauses, as in (41), which should be compared to (36) and (37).

(41) [5'teki uçağa bin-**se**-**ydi**-k] toplantıya yetiş-e-mez-**di**-k.
take-COND-PASTCOP-1PL make.it-PSB-NEGAOR-PASTCOP-1PL

We would not have made it to the meeting (on time) [if we had taken the 5 AM flight].

It is worth noting that, similar to -*(y)DI* in counterfactual main clauses, a conditional clause marked with -*sAydI* does not necessarily pertain to past time, as in (42), a revised version of (39).

(42) [Yapacak çok işim olma**saydı**] ben de yarınki partiye gelir**di**m.

I would have come to the party tomorrow, too, [if I had not had a lot to do].

Although the main clause of a counterfactual conditional is most frequently marked by -*(A/I)rdI* (or its negated version -*mAzdI*), it is also possible for the form -*(y)AcAktI* to be used when reference is being made to a firmly planned action, a scheduled event, or a situation that is perceived for some other reason as certain to have resulted if the condition had been fulfilled, as in (43).

(43) [Tatilde buraya gel-ebil-**se**-**ydi**-niz] sizi bir hafta gezdir-**ecek**-**ti**-k.
come-PSB-COND-3PL-P.COP take.sightseeing-FUT-PASTCOP-1PL

[If you had been able to come here during the holiday], we were going to take you around to sightsee for a week.

In Chapter 4, the examination of -*sAydI* will be expanded upon, and -*DIysA* will be incorporated into the aforementioned analysis. The aim is to provide a comprehensive understanding of the usage and function of these verbal markers in conditional clauses through this integration.

d) -sAymIş [conditional marker -sA + evidential copula -(y)mIş]

The construction -*sAymIş* is a synthesis of the verbal conditional marker -*sA* and the evidential copula -*(y)mIş* (Göksel & Kerslake, 2005). A conditional clause distinguished by the presence of -*sAymIş* is consistently accompanied by a main clause that is also marked by -*(y)mIş*, as in (44), indicating that the use of this construction is inextricably linked to its main clause.

(44) [Hastalığı bir hafta önce teşhis ed-il-**se**-**ymiş**] ölümü
diagnose-PASS-COND-EVIDCOP
engelle-n-ebil-ecek-**miş**.
prevent-PASS-PSB-FUT-EVIDCOP

Apparently, [if their illness had been diagnosed a week earlier], their death would have been prevented.

The ambiguity of tense reference, which is a characteristic feature of all sentences distinguished by the presence of *-(y)mIş*, is also present in these conditionals, as in (45).

(45) Veli [Türkiye'de ol**saymış**] düğününüzü kaçırmaz**mış**.
  a. *Veli says that [if he were in Turkey] he would not miss your wedding.*
  b. *Veli says that [if he had been in Turkey] he would not have missed your wedding.*

It is important to acknowledge that, due to the incompatibility of combining the conditional copula *-(y)sA* and the evidential copula *-(y)mIş* on one stem, only hypothetical and counterfactual predictive conditionals can be evidentially marked (Göksel & Kerslake, 2005). In other forms of conditional sentences, such as the open conditional exemplified in (46), the evidential marking is restricted to the main clause alone. This limitation highlights the unique grammatical characteristics of hypothetical and counterfactual predictive clauses within the broader context of evidential marking in conditional sentences.

(46) Veli [hava kötü olur**sa**] maçta oynamayacak**mış**.
Apparently, Veli is not going to play in the game [if the weather is bad].

Compound constructions with *olsa-* are parallel in terms of their relative tense and aspectual values to their counterparts with *olursa* (see the *olursa* examples 25-29 under 2.2.2a). They can appear in both hypothetical and counterfactual contexts. The counterfactual variants are distinguishable by the past copular marking of the main clause and by the mutual substitutability of *olsa* and *olsaydı* in the conditional clause. The sentences in (47), (48), and (49) serve as examples of combinations with *-mIş*: (47) may be spoken during a before-the-game discussion, whereas (49) could only be spoken after the game had ended. (48) one could be used in either situation, but if spoken before-the-game, it would probably be as an argument against a decision that has already been made, or as in this context, as a comment on someone's prediction about the game. The form *-mIş olsaydı* is frequently employed as a synonym for *-sAydI* to indicate a counterfactual condition, without any implicit meaning of relative tense.

(47) [Maç uzatmaya git-**miş ol-sa**] seyirciler son metroya yetiş-e-**mez**.
go-PF AUX-COND make.it-PSB-NEGAOR
[If the game had gone to extra time], the audience would not be able to catch the last metro.

[Maç uzatmaya git**miş olsa**] seyirciler son metroya yetiş-e-**mez-di**.
make.it-PSB-NEGAOR-PASTCOP

[If the game had gone to extra time], the audience would not have been able to catch the last metro.

[Maç uzatmaya git-**miş ol-sa-ydı**] seyirciler son metroya yetişe**mezdi**.
go-PF AUX-COND-PASTCOP
[If the game had gone to extra time], the audience would not have been able to catch the last metro.

A comparable gradation of semantics is evident in the use of the *-(I)yor olsa* construction, as well as in the relative-future meaning conveyed through the utilization of *-(y)AcAk olsa*[5].

2.2.3. Knowable Conditionals

The realization of the condition articulated within a predictive conditional clause is inherently uncertain, as it pertains to the unpredictable future, a hypothetical scenario, or an event that has not yet occurred. Conversely, the realization of a condition that is considered knowable is one for which information pertaining to its fulfillment or non-fulfillment is potentially accessible, as it pertains to the present or past, or to planned or scheduled future events (Göksel & Kerslake, 2005).

In a knowable condition, the marker employed is always the conditional copula, which is affixed to one of the following grammatical elements: (i) a tense/aspect/modality suffix of *position 3*[6], with the exception of the aorist, (ii) any constituent functioning as a *subject complement*[7], (iii) the particle *değil*, (iv) one of the existential expressions *var* or *yok*, and (v) the past or evidential copula. Examples (48), (49), and (50) represent knowable conditional structures, and they illustrate the utilization of past, present, and future tense, respectively.

(48) [Sucuklar beş dakika kızar-**dı-ysa**] ol-muş-tur artık.
grill-PF-CONDCOP be.done-PF-GM
[If the sucuk (slices) has (been) grilled for five minutes], they will be done by now.

(49) [Veli basketbol oyn-**uyor-sa**] iyileş-miş ol-malı.
play-IMPF-CONDCOP get.better-PF AUX-OBLG
[If Veli is playing basketball], he must have gotten better.

(50) [Uçak 11'de kalk-**acak-sa**] 9'da havaalanında ol-malı-yız.

take.off-FUT-CONDCOP be-OBLG-1PL

---

[5] See Section 27.2.6 of Göksel & Kerslake (2005) for a trio of examples for both forms.
[6] See Section 8.2.3.3 of Göksel & Kerslake (2005) for details on *position 3*.
[7] See Section 12.1.1.2 of Göksel & Kerslake (2005) for details on *subject complements*.

[If the plane is going to take off at 11 PM], we have to be the airport by 9 PM.

In certain instances, knowable conditions do not indicate any awareness on the part of the speaker regarding the actualization of the condition in question. This is particularly evident in cases where the condition pertains to a present-tense scenario, as in (51) (Göksel & Kerslake, 2005). This category of knowable conditionals exhibits a similar relationship to the scenario articulated within the main clause, as in causing or enabling it, as is observed in predictive conditionals.

(51) [Vaktin var**sa**] gel bir kahveciye gidelim ve sohbet edelim.
[If you have time], let's go to a coffee shop and chat.

In the majority of knowable conditions, the relationship between the two clauses is dissimilar, and comparatively more open-ended. The conditional clause conveys information that the speaker has recently obtained, either from a conversational participant or from an external source within the common context of the speech participants. This condition is considered to be more or less actualized, and presented as a provisional assumption that serves as the backdrop for the speaker's statement within the main clause. Certain types of utterances are commonly observed in the main clause of a knowable condition, such as inferences, as exemplified in (49), suggestions, requests, or other volitional utterances, as exemplified in (51), expressions of obligation, as exemplified in (52), or questions, as exemplified in (53). Questions following conditions that are presented as working assumptions are typically reproachful in tone, indicating a discrepancy between the situation assumed in the conditional clause and some observed behavior or action (Göksel & Kerslake, 2005).

(52) [Ali cinayeti Veli'nin işlediğini biliyor**duysa**] polise söylemeliydi.
[If Ali knew it was Veli who committed the murder], he should have told the police.

(53) [Parası yokmuş**sa**] ne maksatla araba almaya kalkmış?
[If (as is claimed) they have no money], why have they tried to buy a car?

It should be noted that clauses marked with *-sA* or *-(y)sA* do not always convey a conditional meaning. Please see Section 27.4 of Göksel & Kerslake (2005) for examples of (i) *-sA…-sA*, which creates a specific type of *concessive* conditional clause that limits the applicability of the main predicate to the constituent immediately preceding it, while concurrently expressing the probability of non-occurrence of even that event, (ii) *-sA* with *dA*, which can mean *even if*, express the possibility or acceptability of an alternative course of action with *olur* (it will be all right), or mark each of two alternative conditions in a predictive conditional sentence, indicating that the outcome will be the same regardless of which one is actualized, and (iii) *-(y)sA* with *dA*, which can be equivalent to *although*.

2.2.4. Universal Conditionals

When a conditional clause incorporates a question phrase, often referred to as a *wh-phrase*, the semantic meaning of the said clause is congruent to that of a clause utilizing the correlative pronouns *whoever*, *wherever*, etc. in English. This particular type of conditional clause is referred to as *universal*, as it encompasses an unrestricted range of conditions [8]. It is noteworthy that within these clauses, the suffix *-sA* can be interchangeably utilized with various verbal *-(y)sA* combinations, as in (54) (Göksel & Kerslake, 2005).

(54)	[**Kim**e	sor-du-k-**sa**/sor-du-**ysa**-k[9]/sor-**sa**-k] aynı cevabı aldık.
	who-DAT	ask-PF-1PL-CONDCOP/ask-PF-CONDCOP-1PL/ask-COND-1PL

[Whoever we asked] we got the same answer.

It is occasionally observed that within subordinate clauses, the wh-phrase is reflected through the utilization of a *resumptive pronoun* in the main clause, as in (55), a revised version of (54).

(55)	[**Kim**e sordu**ysa**k]	**hepsi**-nden aynı cevabı aldık.
		all.of.them-ABL

[Whoever we asked] we got the same answer (from all of them).

In instances where the question phrase employed is *kim* (who), *ne* (what), or *nere-* (where), the concept of universality can be further emphasized by the inclusion of the indefinite pronoun *her* (every), preceding said item, as in (56).

(56) Veli [**her nereye** gider**se**] sevimli köpeğini de yanında götürüyordu.
Veli took his cute dog with him [wherever he went].

It is common to observe that within the framework of universal conditionals, the question phrase *nasıl* (how) is utilized in a manner that is detached from any form of conditionality and instead serves to convey a sense of similarity between two given situations, as in (57). In contrast, when *nasıl* is employed in a manner that serves to modify the verb within the subordinate clause, it is reflected through the utilization of the adverb *öyle* (so/in that way) in the main clause, as in (58).

---

[8] Except in instances where the question phrase employed is *hangi* (which).

[9] *Sorduksa* and *sorduysak* have the same signification; Göksel & Kerslake (2005) do not provide an equivalent to s*orduysak* in their example but it should have been there as it is most probably the more common of the two.

(57) [Veli **nasıl** akşamları taksicilik yapıyor**sa**] sen de çalışmalarına engel olmayacak bir iş bulabilirsin.
[Just as Veli works as taxi driver in the evenings], you too could find a job that would not interfere with your studies.

(58) [Çocuklar babalarını **nasıl** seviyor**sa**] çoğu Galatasaraylı da Fatih Terim'i **öyle** seviyor.
Most Galatasaray fans love Fatih [in the same way that children love their father].

A significant proportion of universal conditional constructions convey a sense of *concession*, particularly with regard to the failure (predicted or actual) to attain a stated objective, despite the efforts indicated within the conditional clause. Please see Section 27.5.1 of Göksel & Kerslake (2005) for examples that demonstrate this meaning using *dA* after conditional verb and through using the aorist + *-(y)sA* structure immediately followed by an imperative or optative form of the same verb in the same grammatical person.

2.2.5. Other Ways to Convey Conditionality/Counterfactuality

Just as there are other ways to convey conditionality and counterfactuality in English in addition to the canonical form of *if X, (then) Y* (see 2.1), there exists different ways of conveying a conditional or counterfactual meaning in Turkish, as in (Göksel & Kerslake, 2005).

(59) Ben bu rengi seç**mezdi**m.
I would not have chosen this color.

(60) Daha uygun bir renk seç**ebilirdi**niz.
You could have chosen a more appropriate color (but you did not).

(61) O filme gitmek hoş ol**urdu**/ol**acaktı**.
It would have been nice to go to that movie.

(62) [Ali Veli'ye yardım ed**ebileceği hâlde**] hiçbir şey yapmadı.
[Although Ali could have helped Veli], he did nothing.

(63) Yerde bulduğun cüzdanı geri ver**meliydi**n.
You should have returned that wallet you found on the floor (but you did not).

(64) O sorunu sabahki toplantıda değerlendirmemiz **lazımdı**/**gerekiyordu**.
We were supposed to discuss that issue at the meeting this morning (but we did not).

(65) **Keşke** Veli'ye söyleme**seydi**n/söyleme**yeydi**n.
If only you had not told Veli.

(66) **Keşke** başka bir ülkede yaş**ıyor olsam**.
If only I were living in a different country.

(67) Bu saatte çalış**ıyor olmalıydı**n.
I should be working at this time of day (but you are not, you are doing something else).

(68) Dosyalar saat 5'e kadar gönderil**miş olurdu**.
By 5 PM, the files would have been sent.

(69) [O ünlü *Altılı Masa* güçlü bir aday çıkarama**dığı takdirde**] seçim kaybedilecektir.
If that famous Table-of-Six fails to announce a strong candidate, the election will be lost.

(70) [Güçlü bir aday çıkarılama**ması durumunda/hâlinde**] seçim kaybedilecektir.
[In the event of a failure to announce a strong candidate], the election will be lost.

(71) [Güçlü bir aday çıkarıl**dı mı]** seçim kazanıldı demektir.
[If/Once a strong candidate is announced], the election is as good as won.

CHAPTER 3

3. ON THE GRAPHICAL REPRESENTATION OF COUNTERFACTUALS

Felix qui potuit rerum cognoscere causas
Atque metus omnes, et inexorabile fatum
Subjecit pedibus, strepitumque Acherontis avari

(He who's been able to learn the causes of things is happy, and has set all fear, and unrelenting fate, and the noise of greedy Acheron, under his feet.)

Virgil - Georgica II, v. 490 (Translation taken from Poetry in Translation)

This chapter focuses on how counterfactuals can be represented graphically. The emphasis is on understanding causal inference through the use of BNs. The chapter introduces the reader to the *causal modeling* part of the approach that is to be introduced in Chapter 4.

3.1. On Causality and Causal Inference

This subchapter will focus on Pearl & Mackenzie (2018), which is a thought-provoking and informative exploration of the concept of causation and its importance in science and philosophy that argues that understanding causation is essential for making sense of the world and for advancing scientific knowledge.

Causal inference is the process of identifying and understanding the relationships between variables, with the goal of determining whether a change in one variable causes a change in another. This is often done in order to understand the underlying causes of observed outcomes and make informed decisions based on this knowledge.

Judea Pearl is a leading researcher in the field of causal inference and has contributed significantly to our understanding of this topic through his work on causal modeling and graphical models. According to Pearl & Mackenzie (2018), the question *why* has long been the driving force behind human progress, enabling us to make remarkable discoveries and innovations throughout history. As a result of our ability to understand cause and effect relationships, we have been able to create complex societies and, ultimately, the advanced civilization in which we currently reside. This unique capability, referred to as *causal inference,* enables the human brain to store and utilize vast amounts of knowledge about the ways in which various factors can influence one another. This

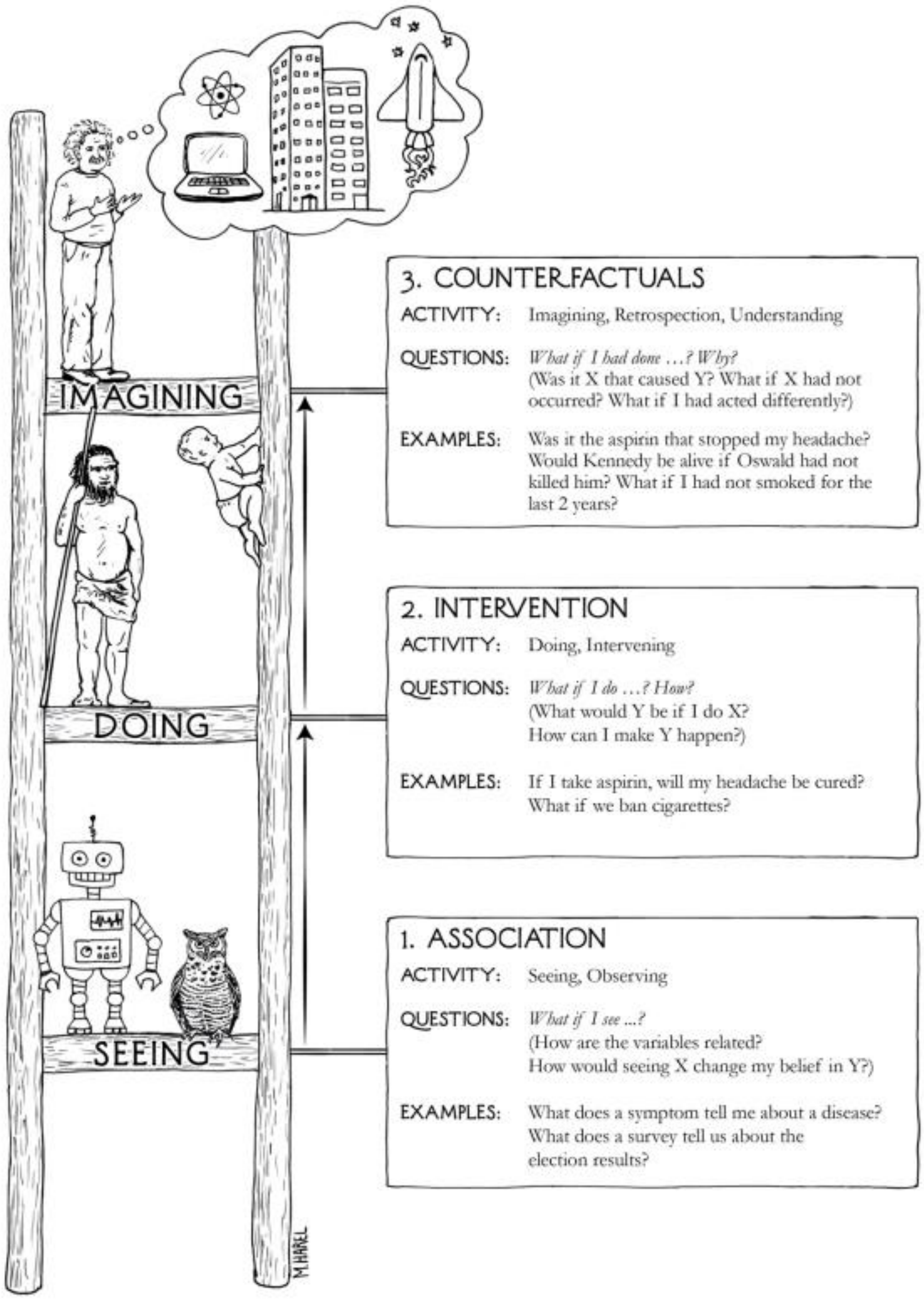

Figure 1: An illustration of the Ladder of Causation by Maayan Harel from Pearl & Mackenzie (2018).

knowledge allows us to pose and answer a plethora of questions concerning a range of topics, from the efficacy of treatments to the causes of health issues and the impact of laws and policies. Pearl & MacKenzie (2018) posit that an understanding of causal inference can not only deepen our comprehension of the ways in which humans understand cause and effect, but also enhance our ability to analyze data and make informed decisions.

One of Pearl's most well-known examples of causal inference is the *Ladder of Causation*, which illustrates the levels of influence that different variables can have on one another (Pearl & Mackenzie, 2018). Figure 1 is an illustration of the Ladder of Causation by Maayan Harel.

3.1.1. The First Rung

Inherently, human beings possess a proclivity for analyzing the environment and forming *associations*. This type of cognition resides on the first rung of the Ladder of Causation.

Interestingly, while humans are biologically predisposed to do so from infancy, the artificial intelligence systems they have created to assist them in their daily routines still struggle to attain this level of understanding. Most animals, as well as artificial intelligence programs, remain confined to this first rung of the ladder; for example, an owl tracks its prey by monitoring its movements and attempting to predict its location in the next moment, but it has no interest in the reason for the prey's movement. Self-driving vehicles, although appearing futuristic, are also limited to the first step of the ladder. As they are programmed to only react to observations, an autonomous car cannot anticipate the various potential reactions of a drunken pedestrian crossing the road in response to a car horn. All possible scenarios would have to be explicitly programmed into the car for it to be able to respond appropriately to each one.

Data collection can also be considered to exist on the first rung of the ladder as it involves projections based on passive observation; for instance, a marketing director may be asked to determine the likelihood of a toothpaste-buying customer also purchasing dental floss. They would likely gather data on the number of toothpaste-buying customers and floss-buying customers. This query can be denoted as $P(\text{floss} \mid \text{toothpaste})$, or *What is the probability of floss, given that you see toothpaste?* These types of questions form the fundamental foundation of statistics; however, they do not provide insight into cause and effect. How can the marketing manager determine whether toothpaste or floss is the cause? While examining sales of dental hygiene products may not be of significant importance, it is evident that simply observing basic probability is insufficient in most other cases.

3.1.2. The Second Rung

Ascending the Ladder of Causation requires not only observing the world, but actively manipulating it through *interventions*. Using this ability regularly is uniquely human.

The second rung of the ladder is defined by the question, *What if we do...?* The key element here is the *do*; this rung is characterized by actively influencing outcomes, as opposed to the passive observations of the first rung. Consider the scenario of taking a painkiller to alleviate a headache, this is a purposeful intervention designed to alleviate the discomfort being experienced.

Returning to our example of the dental hygiene marketing manager, they may ask, *Will floss sales be impacted if we alter the price of toothpaste?* Contrastingly, it is currently not possible to program computers to accurately pose these types of questions, which is why they are unable to progress beyond the first rung of the ladder.

One of the most effective methods of evaluating the effect of a particular factor is through the implementation of a controlled experiment. This involves creating groups that are as similar as possible and applying a test to one group but not the other. This allows for the variable and its impact to be objectively measured in isolation.

Controlled experiments are not a novel concept; they are even mentioned in the Bible. In the story of Daniel, the Babylonian King Nebuchadnezzar captured nobles from Jerusalem and brought them to his court, as was customary. This involved educating them in the elite Babylonian diet of rich meats and wine; however, in accordance with Jewish dietary laws, some of the Jewish boys refused to eat the meat. Daniel was one of them. He proposed that he and three of his friends be given a vegetarian diet, while another group of boys were given the king's diet. This second group would be referred to as a control group in modern terms. After ten days, they compared the results, and Daniel's group was significantly healthier. As a result, Nebuchadnezzar appointed them to high court positions.

A more contemporary example is the use of controlled experiments by companies such as Facebook, which frequently tests different configurations of items on web pages and compares the results of different groups exposed to these various arrangements.

3.1.3. The Third Rung

The third rung of the ladder, exclusive to humans, is the ability to envision how different interventions can lead to divergent outcomes. One method of applying this imaginative capacity is through the utilization of counterfactual models, or visualizing what would occur if a different action were taken; for instance, climate scientists frequently employ this approach by posing queries such as, *Would we experience extreme heatwaves if atmospheric carbon dioxide levels were at preindustrial levels?*

Counterfactuals can also be applied to past events. They are commonly employed in legal proceedings in the form of *but-for* causation questions; for example, in a trial addressing a case where an individual has been shot and killed, the central question is, *But for the defendant pulling the trigger, would the victim have died?* Machines are incapable of comprehending these types of counterfactual questions. If a house were to catch fire after someone struck a match, it is natural for most people to assert that the house would still be standing had the match not been lit; however, it is also logically true that the house would remain intact if oxygen were not present (see also Mackie, 1965). While oxygen is a normal and expected presence, lighting a match is not, so we disregard the causal relationship between oxygen and the fire. A computer does not think in this manner. For a computer, both the lit match and oxygen would be considered equal factors. In mathematical terms, both are *necessary causes*. As a result, the computer may be just as likely to conclude that oxygen caused the fire. A computer may also calculate whether the match was a *sufficient cause* of the fire. This means that although other factors may have been necessary for the fire to start, the computer determines whether the match was sufficiently responsible to be considered the cause. If it had been programmed to recognize that oxygen was necessary for fire, it may conclude that oxygen was the cause.

All in all, it is vital to comprehend the three levels of causation in order to effectively analyze causal inquiries.

3.2. Graphical Representation Through Causal Diagrams

This subchapter will make heavy use of Pearl (2009) and Pearl et al. (2016) to introduce the graphical notation and terminology required to understand the diagrams used for causal inferences.

3.2.1. Directed Acyclic Graphs (DAGs)

The colloquial usage of the term *graph* often refers to a broad range of visual aids, often interchangeably with *chart*; however, in the realm of mathematics, a graph is a formally defined object. Figure 2[10] depicts a mathematical graph which comprises a set of *nodes* represented by *A*, *B*, and *C*, (the circles) and a set of *edges* (the lines) connecting these nodes represented by points *X* and *Y*. The nodes within a graph may be connected or unconnected by the edges.

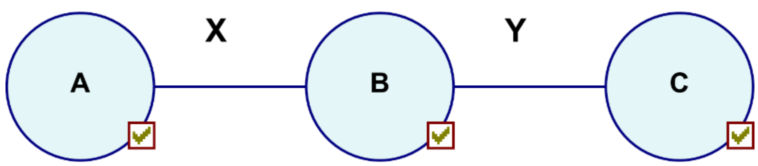

Figure 2: An undirected graph in which nodes are named A, B, and C, and edges are named X and Y.

In a graph, two nodes are said to be *adjacent* if there exists an edge connecting them. As depicted in Figure 2, nodes A and B, as well as B and C, are adjacent to one another,

[10] In this work, figures that involve graphical representations of causal relations were created using the GeNIe Academic Version (BayesFusion, 2022). All the .xdsl files will be made available online at https://bit.ly/ilteristhesis.

whereas A and C are not. A graph in which an edge connects every pair of nodes is known as a *complete* graph.

A sequence of nodes connecting node A to node B, such that each consecutive pair of nodes is connected by an edge, is referred to as a *path* between nodes X and Y. In Figure 2for instance, a path exists between nodes A and C as A is connected to B via an edge and B is connected to C via an edge.

Edges in a graph can be classified as either *directed* or *undirected*. The edges in Figure 2 are undirected as they do not possess a designated *in* end or *out* end. In contrast, a directed edge connects one node to another in a specific direction, as indicated by an arrowhead. When all of the edges in a graph are directed, the graph is referred to as a *directed* graph. An example of a directed graph is illustrated in Figure 3, where edge X is directed from node A to node B and edge Y is directed from node B to node C.

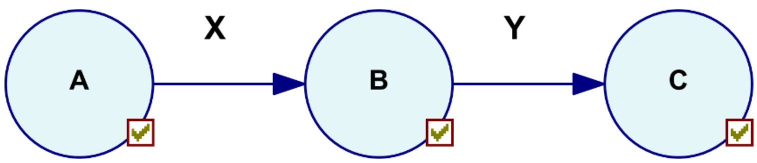

Figure 3: A directed graph in which node A is a parent of B and B is a parent of C.

In a directed graph, the node from which a directed edge originates is referred to as the *parent* of the node to which the edge leads, while the latter node is considered the *child* of the former. Using the graph in Figure 3 as an example, node A is the parent of node B, and node B is the parent of node C; correspondingly, node B is the child of node A and node C is the child of node B. A path between two nodes in a directed graph is considered a directed path if it can be traced along the arrows, without encountering any nodes that have two edges directed into them or two edges directed out of them. If two nodes are connected by a directed path, the first node can be considered the *ancestor* of all nodes on the path, and all nodes on the path can be considered the *descendants* of the first node. In other words, the relationship between ancestor and descendant nodes in a directed graph can be analogized to the relationship between parent and child nodes; for instance, in Figure 3, node A is the ancestor of nodes B and C, and both nodes B and C are descendants of node A.

A directed path that originates and terminates at the same node results in a cyclic path and, subsequently, a *cyclic* graph. In contrast, a directed graph that lacks any cyclic paths is considered *acyclic*; for example, as shown in Figure 4, the graph on the left is cyclic due to the existence of directed paths from node X back to itself, while the graph on the right is acyclic, as there is no directed path from any node to itself.

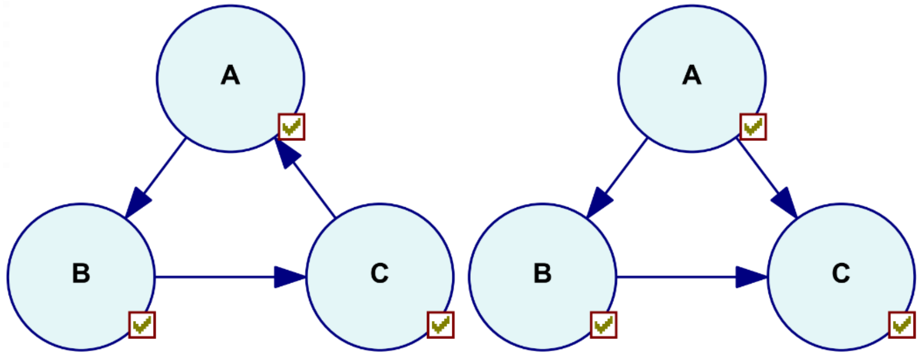

Figure 4: A cyclic graph on the left and an acyclic one on the right.

3.2.2. Conditional Independence

Directed acyclic graphs (DAGs) are often used in causal modeling to represent the relationships between different variables in a system. In a causal model, the edges in the DAG represent the direct causal relationships between variables, with the direction of the edge indicating the direction of the causal influence; for example, if variable A is believed to directly cause variable B, then there would be a directed edge from A to B in the DAG.

One important concept related to DAGs is *conditional independence*. This refers to the idea that observing the value of one variable does not affect the probability distribution of another variable; for example, if variables A and B are conditionally independent given variable C, then the probability of A given B and C is the same as the probability of A given C alone. In a DAG, conditional independence can be represented by the absence of an edge between two variables.

There are three important node structures to consider in a DAG: chains, forks, and colliders, which are depicted in Figure 5.

A *chain* refers to a sequence of three or more nodes where there is a directed edge connecting each consecutive pair of nodes; for example, in the DAG on the left in Figure 5, nodes A, B, and C form a chain because there is a directed edge from A to B, and a directed edge from B to C. In a chain, when there exists only a single unidirectional path between variables A and C, and the presence of a set of variables, denoted as B, results in the interception of that path, then A and C are considered conditionally independent given B.

A *fork*, also known as a *diverging* structure, refers to a situation where a single node is connected to two or more other nodes by directed edges; for example, in the DAG in the middle in Figure 5, node B is connected to nodes A and C by directed edges, forming a fork. In a fork, when a common cause, denoted by B, exists for variables A and C, and there is only a single path linking A and C, then A and C are considered conditionally independent given B.

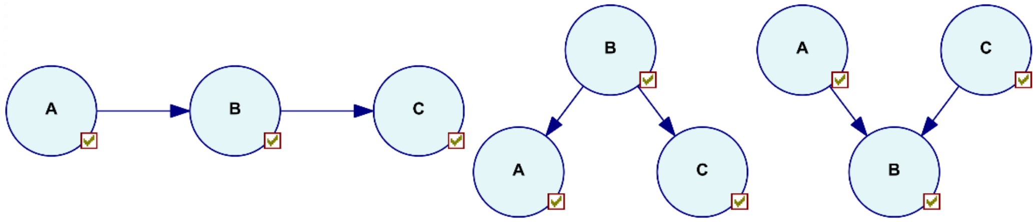

Figure 5: A chain on the left, a fork in the middle, and a collider on the right.

A *collider*, also known as a *converging* structure, refers to a situation where two or more nodes are connected to a single node by directed edges; for example, in the DAG on the right in Figure 5, nodes A and C are both connected to node B by directed edges, forming a collider. In a collider, when a collision node, denoted by B, exists between variables A and C and there is only a single path connecting them, then A and C are unconditionally independent but become dependent when conditioned on B and any descendant variables of B.

3.2.3. d-separation

It is uncommon for a graphical causal model to comprise a single path linking variables; instead, numerous potential paths connecting pairs of variables, comprising various chains, forks, and colliders, are typically present in the majority of graphical models.

Directional separation, more commonly known as *d-separation*, is a process built upon the conditional independence rules explained above that is used to determine dependencies shared by all data sets generated by a graphical causal model of any complexity. If all paths between two nodes A and B are blocked, they are considered *d-separated*, which indicates a definitive independence between the variables they represent. If even one path between A and B is unblocked, they are considered *d-connected*, which signifies a potential or likely dependence between the variables they represent. In this sense, Pearl et al. (2016) liken the paths between the variables to how water flows through piping: if one or more pipes is unblocked, water can flow between the variables at either end; however, if a single pipe is blocked, it can prevent the flow of water through all paths.

The ability of certain nodes to block paths during *d*-separation depends on whether the analysis being performed is unconditional or conditional. When no variables are being conditioned on, only colliders are capable of blocking paths. This is due to the fact that unconditional dependence cannot pass through colliders, as previously mentioned. Therefore, if every path between two nodes A and C contains a collider, A and C must be marginally independent, as they cannot be unconditionally dependent; on the other hand, when a set of variables X is being conditioned on, both colliders and certain noncolliders can block paths. Specifically, a collider not included in X, and without any descendants in X, can block a path, as well as a chain or fork with a middle node in X. This is because, as previously discussed, conditioning on these types of nodes can lead to the independence of the variables at the ends of the paths. So, to give a formal definition of *d*-separation (see also Definition 2.4.1 from Pearl et al, 2016):

A path *p* is blocked by a set of nodes *X*, if and only if

- *p* includes a chain of nodes A → B → C, or a fork with nodes A ← B → C, and the middle node B is a member of *X* (in other words, B is conditioned upon), or
- *p* includes a collider with nodes A → B ← C, and neither the collision node B nor any of its descendants are members of *X*.

3.2.4. The "Bayesian" Way: The Bayes Theorem and Bayesian Networks

Bayesian networks (BNs), introduced by Pearl in the 80s, are graphical models that represent the probabilistic dependencies between a group of variables. They are concise representations of joint probability distributions by using conditional probabilities. A BN comprises a qualitative part, DAG, and a quantitative part, conditional probability tables (CPTs). We have already introduced DAGs and we will be making use of CPTs in the examples in the following sections; so, this section will focus on the basis of BNs.

The theoretical foundation of BNs can be traced back to the seminal work of Reverend Thomas Bayes. Bayes' *essay towards solving a problem in the doctrine of chances* was posthumously published by Richard Price in 1763, and it presented a formulation of what is now known as the *Bayes' theorem*, which pertains to the computation of conditional probabilities. This pivotal work has served as a significant contribution to the field of statistics and probability theory, inspiring further developments in Bayesian inference and modeling.

Bayes' theorem is a fundamental concept in probability theory that illustrates the relationship between conditional probabilities. It asserts that the probability of an event X, given the occurrence of event Y, can be computed as the product of the inverse probability of event Y given event X and the probability of event X, divided by the probability of event Y. Mathematically, it can be expressed as:

$$P(\mathrm{X} \mid \mathrm{Y}) = \frac{P(\mathrm{Y} \mid \mathrm{X})P(\mathrm{X})}{P(\mathrm{Y})}$$

At first glance, Bayes' theorem may appears to have limited practical value, as it entails the calculation of a single probability term *P*(Y | X) in terms of three other probability terms: *P*(X | Y), *P*(Y), and *P*(X); however, the rule can be useful in practice as it enables the computation of the fourth probability term, which is often needed when we have reliable probability estimates for the three aforementioned terms. In many instances, we observe the effect of an unknown cause and seek to infer the cause itself. In such cases, the theorem can be expressed as (Russell & Norvig, 2021):

$$P(\text{cause} \mid \text{effect}) = \frac{P(\text{effect} \mid \text{cause})P(\text{cause})}{P(\text{effect})}$$

The conditional probability *P*(effect | cause) denotes the *causal* direction, whereas *P*(cause | effect) refers to the *diagnostic* direction. In medical diagnosis, for example, there often exist conditional probabilities on causal relationships; a doctor, for instance, may know *P*(symptoms | disease) in general and wish to compute a diagnosis, *P*(disease | symptoms) for specific patients. Consider a scenario where a doctor is aware that meningitis causes a stiff neck in 75% of cases. Additionally, the doctor has unconditional information, such as the prior probability of a patient having meningitis being 1/40000 and the prior probability of a stiff neck being 0.5%. Using propositions S and M to represent a stiff neck and meningitis, respectively, we can compute *P*(M | S) with (adapted from Russell & Norvig, 2021):

$$P(\mathrm{M} \mid \mathrm{S}) = \frac{P(\mathrm{S} \mid \mathrm{M})P(\mathrm{M})}{P(S)} = \frac{0.75 \times \frac{1}{40000}}{0.005} = 0.00375$$

This indicates that 0.375% of patients with a stiff neck are likely to have meningitis. Despite the high correlation between meningitis and a stiff neck, the prior probability of stiff necks from any cause is much greater than that of meningitis.

We can avoid the need to assess the prior probability of evidence, *P*(S) through computing the posterior probability for each value of the query variable, M and ¬M, and normalizing the results. The normalization of Bayes' rule can be represented as follows, where the normalization constant α ensures that the probabilities in the resulting distribution sum to 1 (Russell & Norvig, 2021):

$$P(\mathrm{Y} \mid \mathrm{X}) = \propto P(\mathrm{X} \mid \mathrm{Y})P(\mathrm{Y})$$

The availability of conditional probability in one direction but not the other raises the question of why this is the case. In the case of meningitis, a doctor may know that a stiff neck implies meningitis in 1 out of 3000 cases, thus having diagnostic information from symptoms to causes; however, diagnostic knowledge is often less reliable than causal knowledge. For example, if there is a sudden epidemic of meningitis, the unconditional probability of meningitis, *P*(M), will increase. A doctor who derived the diagnostic probability *P*(M | S) directly from statistical observation of patients before the epidemic will have no means to update the value; in contrast, the doctor who computes *P*(M | S) using Bayes' rule and the other three values will recognize that *P*(M | S) should increase proportionally with *P*(M). Crucially, causal information, such as *P*(S | M), remains unaffected by the epidemic since it reflects the way meningitis operates. The use of direct causal or model-based knowledge is necessary for ensuring the robustness of probabilistic systems in practical settings (Russell & Norvig, 2021).

While Bayes' theorem is generally applicable to all probability distributions, its computation becomes progressively more demanding when dealing with more complex problems than the one discussed above. The factorization of joint probability distributions over a set of *n* random variables is known to have *n*! equivalent forms, each expressing the distribution as a product of conditional probability distributions for each variable given the other variables; however BNs can effectively model complex systems with numerous variables because, as we mentioned before, they represent joint probability distributions in a compact and intuitive manner. In other words, BNs make it feasible to compute Bayesian inference on large and interrelated probability distributions.

BNs are represented as directed acyclic graphs. In DAGs, each node represents a variable and each directed edge represents a probabilistic relationship between variables. The strength of the probabilistic relationships between variables in a BN is specified using CPTs. Each node in the graph has a CPT associated with it, which describes the probability of that variable given the values of its parent variables. A graphical representation of the aforementioned meningitis example can be seen in Figure 6[11].

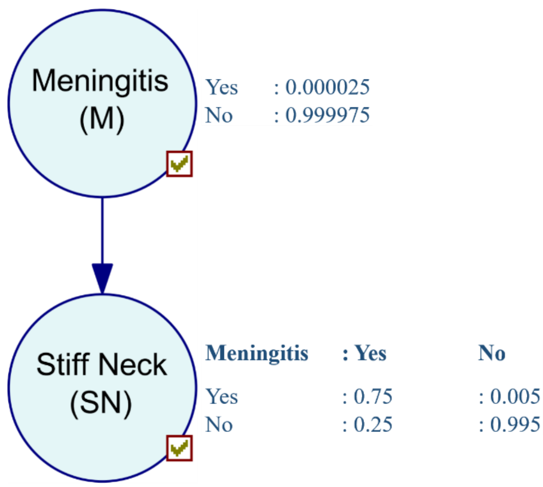

Figure 6: The BN that represents the meningitis example.

BNs allow for probabilistic inference. Given a set of observations or evidence, we can calculate the posterior probabilities of other variables in the network using Bayes' rule. An example that represents the calculation of the P(M | S) value from the meningitis example can be found in Figure 7. Note that efficient algorithms are available to compute Bayesian reasoning on BNs, and these algorithms are readily implemented in a variety of BN software (Lauritzen & Spiegelhalter, 1988). When an observation about a variable is instantiated in a BN, these algorithms can be used to compute the posterior probability of the other unobserved variable, as in Figure 7. Throughout this thesis, we use GeNIe Academic Version (BayesFusion, 2022) and the probabilistic reasoning algorithms implemented within this software.

In Figure 7, the BN on the left is just the bar chart view of the BN in Figure 6. The BN in the middle represents the P(S | M) scenario, and is a *causal reasoning* example: observing

[11] For figures with CPTs, the CPTs were added manually using the draw.io / diagrams.net desktop application (JGraph, 2023).

Meningitis (M) increases the chance of a Stiff Neck (SN) from 0.05% to 75%, in other words, a change in the *cause* leads to a change in the *effect*. The BN on the right represents the *P*(M | S) scenario, and is a *diagnostic reasoning* example: setting SN to *yes* increases the probability of M from 0.000025 to 0.00375[12], in other words, a change in the *effect* leads to a change in the *cause*. In this case, however, as we mentioned earlier, the probability of stiff necks caused by any source, 0.005, is significantly higher than the probability of meningitis.

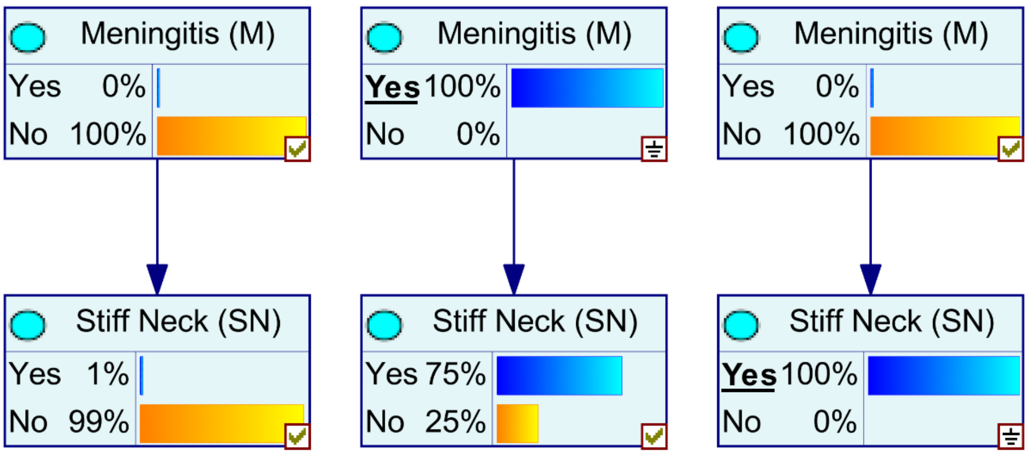

Figure 7: GeNIe bar chart view of the BN in Figure 6 on the left; the probability of meningitis given the patient has a stiff neck on the right.

The following subchapter will help to elucidate these inference-making methods through two different examples.

3.3. Linking a Scenario to a BN

In this subchapter, the aforementioned concepts of the ladder of causation, chains, forks, colliders, and *d*-separation will be exemplified and substantiated through representing two scenarios, two contexts, with BNs. The first one is Pearl's Firing Squad example (Pearl & Mackenzie, 2018), and the second is an original example that focuses on a Cognitive Science MSc student at METU taking COGS515: Artificial Intelligence for Cognitive Science during the COVID-19 pandemic.

3.3.1. A Simple BN Example: The Firing Squad

Imagine a scenario in which a prisoner is about to be subjected to capital punishment by a firing squad. A series of events must transpire in order for this to occur. Initially, the

---

[12] [12] The value of M in the BN on the right in Figure 7 is 0.00375, the 0 shown is a simple GeNIe limitation; the exact value can be seen through Node Properties → Value.

court orders the execution. This order is subsequently conveyed to a captain, who signals the soldiers A and B comprising the firing squad to shoot. It is assumed that these soldiers are compliant and capable marksmen, and will only fire their weapons upon receiving the command. If either soldier A or B fires, the prisoner will die.

A graphical representation of this example can be seen in the diagram on the left in Figure 8. Each node represents a binary variable; for example, if CO is *issued* then the court has ordered the execution and if D is *alive*, then the prisoner is alive. By tracing the connections between variables depicted, we can answer questions from different rungs of the ladder of causation.

As previously mentioned, the first rung is about associations, and it is where we can address questions such as *If the prisoner is dead, was the court order issued?* We can follow the directional connections depicted in the BN and use standard logic to see that the answer to the question must be *yes*: the chain of events leading to the execution requires the soldiers to fire, so, the captain must have signaled them, and for the captain to have done so, they must have received the court order. A graphical representation of this example can be seen in the diagram on the right in Figure 8.

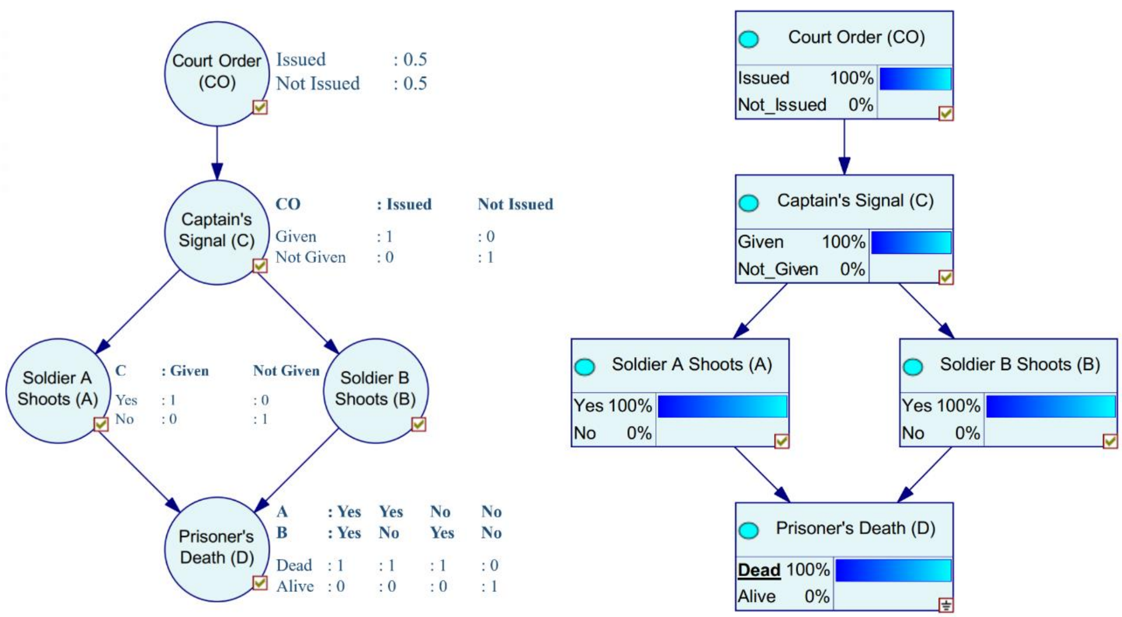

Figure 8: The BN that represents the Firing Squad example from Pearl & Mackenzie (2018) on the left (GeNIe icon view); an *observation* example which involves the prisoner's death on the right (GeNIe bar chart view).

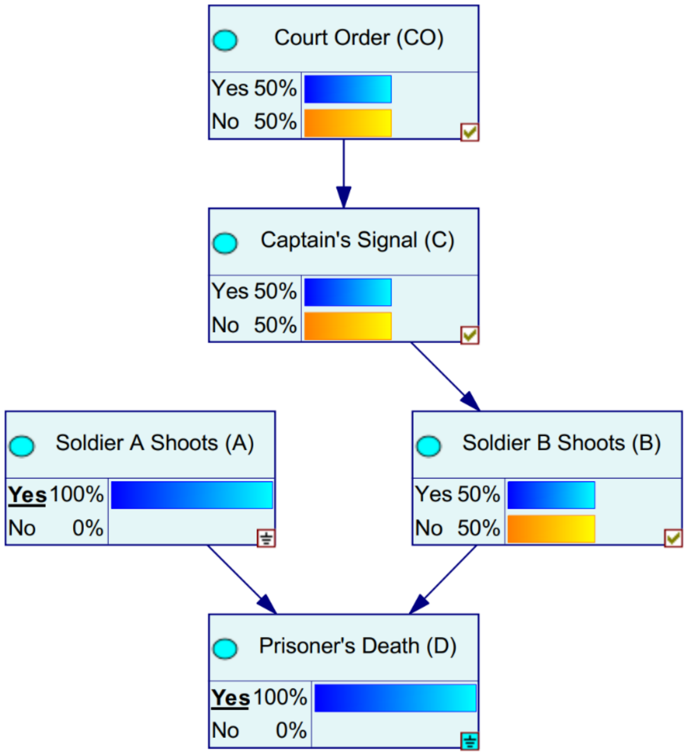

Figure 9: An *intervention* example. Soldier A *decides to* fire, which can be represented through removing the edge from C to A and assigning A the value *yes*.

The second rung is about interventions, and it is where we can address questions such as *What happens to the prisoner if Soldier A decides on his own initiative to fire (without waiting for the captain's command)?* The question is somewhat paradoxical in nature, in that it was previously stated that the soldiers were to shoot only when the captain signals. Therefore, to depict this on a diagram, we must *intervene*, that is, remove all lines of causality that influence A and assign it the value *yes*, as in Figure 9. This operation of removing incoming edges and instantiating a variable is also known as *graph surgery* and is used to model interventions in BNs (see *do*-calculus from Pearl, 2009 and Shpitser & Pearl, 2008). This implies that the prisoner will be dead because A will have shot.

The distinction between seeing/association/rung 1 and doing/intervening/rung 2 may be discerned through comparing the outcomes for Soldier B in the two questions: If we *witness* A firing their weapon, it implies that B also discharged their weapon; however, if A *chooses to* fire, or if we *force* them to fire, we would be uncertain as to whether B fired or not since it remains contingent on the captain's signal.

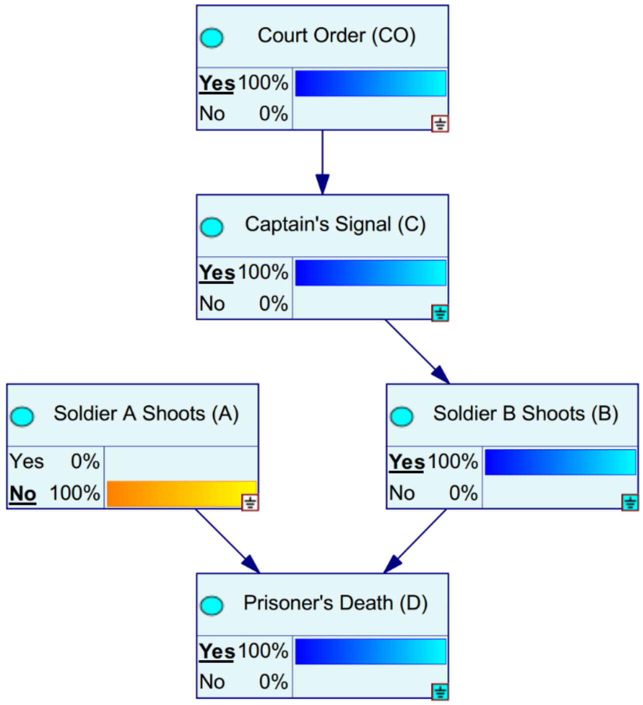

Figure 10: A *counterfactual* example. The prisoner is dead, but what would be the case if Soldier A had chosen not to discharge their weapon?

The third rung is about counterfactuals, and it is where we can address questions such as *What would have happened to the prisoner if Soldier A had decided not to fire?* Suppose we observe that the prisoner has died, which means that A and B shot upon receiving the captain's signal, who was in possession of the court order. In trying to answer the question, one must compare the actual world with an imaginary, yet contradictory, scenario in which A does not fire, but all other conditions remain unchanged. This process involves eliminating all lines of causality that influence A, as in the intervention example, but this time assigning it the value *false* and maintaining the values of the other nodes as they are in the real world, as depicted in Figure 10.

Of course, notwithstanding the fact that A had altered their decision, the prisoner would still have died in this hypothetical world, as B would still have fired their weapon. According to Pearl & Mackenzie (2018), this is likely one of the reasons why such executions are carried out by a firing squad rather than a single soldier; it ensures that the issued order is fulfilled while distributing a rather heavy burden among those who are responsible.

There are five different variables in this simple Firing Squad example from Pearl and Mackenzie (2018), and as previously mentioned, the values of the nodes are all binary (*yes*/*no* / *dead*/*alive*). Causal reasoning in real life is rarely this straightforward, and some scenarios in which it would be considerably more difficult are as follows:

What if

- there were additional variables?
- the variables were not binary?
- the predictions were quantitative (e.g., forecasting the impact of a minimum wage on unemployment rather than a prisoner's binary state of being alive or dead)?
- there were uncertainties (e.g., a firearm malfunction for the soldier/s, a lapse in judgment for the captain, a heart attack for the prisoner)?

There will be answers to some of these *what if* scenarios in the next example.

3.3.2. A More Complex BN Example: The COGS MSc Student at METU

This example is about a Cognitive Science MSc student at METU taking COGS515: Artificial Intelligence for Cognitive Science during the COVID-19 pandemic. It is much more complex than the Firing Squad; for starters, in that it has double the number of nodes, non-binary variables, and more complex CPTs.

The BN that represents this example consists of ten nodes and the following are the explanations of its nodes:

- *Global Pandemic:* The probability of a global pandemic for any given semester; if this is a one-in-a-hundred-year event, let us say, with 0.005 chance there will be a pandemic in one of the two semesters in a year.
- *Mental Health*: The mental health state of a person (the student/the instructor); stable or unstable. Immediately depends on *Global Pandemic*.
- *Teacher Enthusiasm:* The instructor's overall morale and willingness to teach; high, medium, or low. Immediately depends on *Mental Health*.
- *Motivation*: The student's overall morale and motivation to learn/study; high, medium, or low. Immediately depends on *Mental Health*.
- *Background*: The student's major before COGS; engineering, linguistics, or other. The idea behind this tripartite classification was that the engineering majors have an advantage over others regarding this course, and the creator of the diagram, the author, was a linguistics major.
- *COGS501*: The probability of the student having taken COGS501: Algorithmic Structures in Cognition before COGS515; taken or not taken. COGS502: Symbols and Programming is mandatory for all students but there have been students accepted to the "with thesis" program without taking COGS501, and it is highly probable that having taken COGS501 (and COGS502, of course) affects COGS515 performance.
- *COGS515 Difficulty*: How difficult will COGS515 be for the student; hard, normal, or easy. Immediately depends on *Background* and *COGS501*.
- *COGS515 Grade*: Which grade will the student get from COGS515; good, average, or bad. Immediately depends on *Motivation* and *COGS515 Difficulty*.

- *Overall Interest in AI*: How interested will the student be after taking COGS515; high, medium, or low. Immediately depends on *Teacher Enthusiasm* and *COGS515 Grade*.
- *Thesis Subject*: Will the student decide on a thesis subject related to AI or not; related or unrelated. Immediately depends on *Overall Interest in AI*.

The relationship between the nodes, the variables they represent, and the CPTs for the variables can be seen in Figure 11. The probabilities for the variables are not random at all but not entirely reliable either; they can be called *educated guesses* of the author.

We can, in this model, apply the rules of *d*-separation to point out some conditional independencies:

- Global Pandemic and Thesis Subject are conditionally independent given COGS515 Grade and Teacher Enthusiasm.
- COGS501 and Teacher Enthusiasm are conditionally independent given Mental Health and COGS515 Grade and Thesis Subject.
- Background and Overall Interest in AI are conditionally independent given Mental Health and COGS515 Grade.
- Teacher Enthusiasm and Background are conditionally independent given Global Pandemic and COGS501.
- Thesis Subject and COGS501 are conditionally independent given Mental Health, Global Pandemic and Grade.

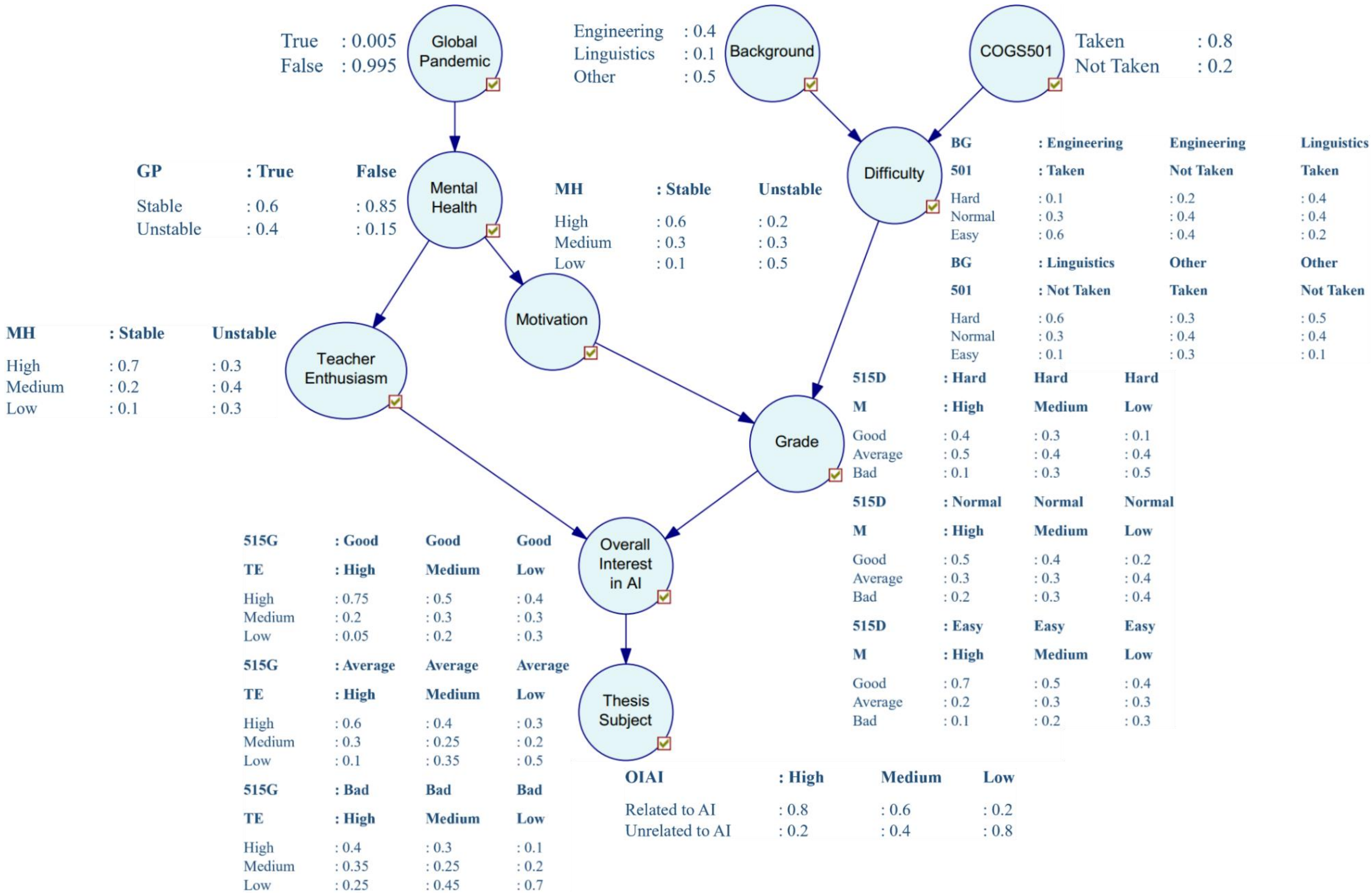

Figure 11: The BN that represents a Cognitive Science MSc student at METU taking COGS515 during the COVID-19 pandemic.

We can build the model in GeNIe[13] (BayesFusion, 2022) to find out the answers to questions such as:

1. What is the probability of a student deciding on a thesis subject related to AI given they are a Linguistics graduate who have taken COGS501 during a global pandemic?

Here, we simply need to *observe*. When we set the Global Pandemic to *true*, Background to *linguistics*, and COGS501 to *taken*, we get 0.6037, which is a plausible result taking into account the initial probabilities. Of course, we can observe additional variables to alter the probability; for example, the probability increases to 0.6358 when Motivation is set to *high*, decreases slightly to 0.6009 when it is set to *medium*, and decreases even further to 0.4612 when it is set to *low*.

2. What is the probability of a student deciding on a thesis subject related to AI given they have an instructor whose enthusiasm does not get affected by a global pandemic, and in turn, its mental health burden?

Here, we need to *intervene*. We need to remove the edge that connects Mental Health to Teacher Enthusiasm and set the latter to one of the values. When we modify the model accordingly, the values for Teacher Enthusiasm variable do not form a CPT, as it is now a discrete distribution[14]. The default probability for a student deciding on a thesis subject related to AI is 0.6281 when the model is as it is in Figure 11, the probability increases to 0.6814 when Teacher Enthusiasm is *high*, decreases to 0.5638 when it is set to *medium*, and decreases even further to 0.4801 when it is set to *low*.

3. What is the probability of a student deciding on a thesis subject related to AI given they are a Philosophy graduate whose motivation was low, who found COGS515 hard but had taken it from a highly enthusiastic instructor during a global pandemic? What would have happened if they had an additional AI background that rendered COGS515 easy for them?

For the first part of the question, we need to *observe*. When we set Background to *other*, COGS515 Difficulty to *hard*, Teacher Enthusiasm to *high*, and Motivation to *low*, we get 0.635. For the second part, we need to *intervene* but keep what we observed for the first part the same. Modifying the model accordingly, namely, eliminating the edges that lead to COGS515 Difficulty and assigning it the value *easy* while keeping the other variables' values the same, we get an increase to 0.67.

[13] We can also build the model in Python using *pomegranate* (Schreiber, 2018), a machine learning package for probabilistic modeling (example code online at https://bit.ly/ilteristhesis).

[14] Let us say we are left with the values under the Stable column: 0.7, 0.2, 0.1. Since we are intervening here, these probability values do not matter, as they are only there to qualitatively tell us which value we should deal with for the Overall Interest in AI variable. They could have been 0.34, 0.33, 0.33 and we would still get the same results.

CHAPTER 4

4. THE PROPOSED APPROACH

Leonard: What the hell are you guys playing?
Sheldon: It's a game we invented. It's called Counterfactuals.
Amy: We postulate an alternate world that differs from ours in one key aspect and then pose questions to each other.
Sheldon: It's fun for ages eight to eighty. Join us.
Leonard: All right. I like a good brainteaser. Give it a whirl.
Sheldon: You're in luck, this is an easy one. In a world where mankind is ruled by a giant intelligent beaver, what food is no longer consumed?

The Big Bang Theory S04E03: The Zazzy Substitution (2010)

In this chapter, the linguistics and causal modeling perspectives introduced in Chapters 2 and 3 are combined through the introduction of a new approach. This approach is then demonstrated through example use cases which highlight the role of context in understanding counterfactual statements.

In Chapter 2, we have discussed that the canonical conditional construction in Turkish involves the use of either the verbal suffix *-sA*, or the copular marker -*(y)*sA in the antecedent and that these two markers were distinguished as: i) *conditional* and *suppositional* by Deny (1921), *counterfactual* and *factive conditional* by Lees (1962), and *subjunctive* and *indicative* by Kuruoğlu (1986) and Sezer (2001) (Deny, 1921; Lees, 1962; Kuruoğlu, 1986 as cited in Sezer, 2001).

The suffix -*DI* and the copular marker -*(y)DI* function as markers of past tense in Turkish. Additionally, the suffix -*DI* also expresses the *perfective* aspect, while the copular marker -*(y)DI* denotes imperfective aspect. The realization of these markers varies depending on their placement in relation to the conditional marker. Specifically, the marker is realized as -*DI* when it precedes the conditional marker, and -*(y)DI* when it succeeds the conditional marker (see Göksel & Kerslake, 2005 and Sezer, 2001 for more details).

The classification of conditional statements as either counterfactual or hypothetical is a matter of ongoing debate among scholars. Ruhi et al. (2000) proposed that the version with the conditional marker at the beginning, *-sAydI*, denotes a counterfactual statement, while its positioning at the end, *-DIysA*, denotes a hypothetical statement. The works of Can Bakırlı (2010) and Oktar & Can (2016) follow the same analysis. Conversely, but not too dissimilarly, Yarbay Duman et al. (2016) call the former as *counterfactual* and the latter as *factual* in their study.

In the remainder of this chapter, the modeling approach we will be using to analyze the difference between *-DIysA* vs *-sAydI* will be introduced through a new example (4.1). Then, that approach will be used to reanalyze the Firing Squad example from earlier (4.2), and later, will be applied to another example, Finding the Medicine (4.3). Penultimately, our approach (**Error! Reference source not found.**) that touches upon the relationship b etween counterfactuals and contextualization, representation, uncertainty, and exogenous factors will be introduced. Finally, the results will be presented (4.5).

4.1. The Party

We had a *what if* section for the Firing Squad example in Chapter 3 (3.3.1) in which we claimed that causal reasoning in real life is rarely that straightforward. One of those scenarios emphasized the fact that uncertainties often make causal reasoning challenging. In this context, we explore another example, the party example from Balke & Pearl (1994), that provides insights into dealing with uncertainties while introducing a different notation to represent counterfactuals.

Imagine a scenario in which Esra, Ali, and Veli belong to a sizeable group of friends who regularly hold parties. Esra has a tendency to attend those parties, while Ali, who has an affection towards Esra, does not share the same interest in social events; therefore, Ali only attends parties when Esra is present, except for rare circumstances. On the other hand, Veli tries to avoid interacting with Esra following their recent breakup, but he has a fondness for parties; consequently, Veli attends parties only when Esra is not present, with rare exceptions. Moreover, Ali and Veli hold a certain animosity towards each other, which often leads to, let us say, conflicts when they encounter each other.

A graphical representation of a party scenario can be seen in Figure 12. Each node represents a binary variable; for example, if E is true, then Esra is at the party, if F is false, then there is not a fight between Ali and Veli at the party.

Let us now consider a conversation between two other friends from the same group, Beril and Kerem, who were not in attendance at the party but were contacted by Ali from his home. The conversation goes as follows:

> Beril: It is likely that Esra is not present at the party; otherwise, Ali would have been there instead of being at home.

Kerem: This implies that Veli is probably present at the party!

Beril: If Ali were indeed present at the party, it is highly probable that he and Veli would engage in a fight.

Kerem: I beg to differ. If Ali were present, Veli would not have been present at the party because Esra would have been there instead.

Beril: This is valid; however, if Ali were present despite Esra's absence, then it is likely that Ali and Veli would be fighting.

Kerem: I concur. It is fortunate that Esra was not there to witness the altercation.

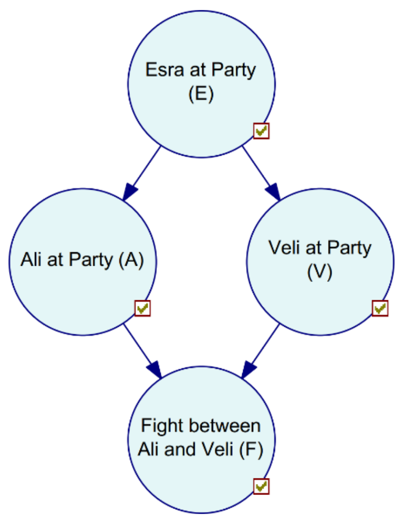

Figure 12: The BN that represents the party example from Balke & Pearl (1994).

In the fourth sentence, Kerem attempts to *explain away* Beril's conclusion by asserting that Ali's being at the party would serve as evidence that Esra was not present, which in turn would imply that Veli was there; however, Kerem approaches Beril's counterfactual statement as an indicative one, assuming that she had directly observed Ali's presence at the party, which enables Kerem to utilize the observation for abductive reasoning. But then again, Beril's subjunctive (counterfactual) statement ought to be interpreted as maintaining everything in the past as it was, including the deductions obtained from abductive reasoning based on actual observations, while simultaneously imposing variables to their counterfactual values.

This example showcases the feasibility of construing the counterfactual statement as a consequence of an external force compelling Ali to attend the party, regardless of any preceding circumstances. We can anticipate that the only variables influenced by the counterfactual supposition would be the descendants of the counterfactual variable, implying that the counterfactual assumption of Ali's attendance does not alter the conviction in Esra's presence from the belief prompted by actual observation.

Now, consider the party example with only E and A in mind, denoting the presence of Esra and Ali, respectively. Suppose that the prior behavior reveals that the probability of Ali being there when Esra is there is 90%, and the probability of Ali not being there when Esra is not there is also 90%. Let us say we observe that both Esra and Ali are absent from the party, would Ali have been there if Esra were there? The answer to this question is dependent on the mechanism underlying the 10% exception to Ali's attendance. If Ali's occasional absence when Esra attends is due to circumstances such as sickness or work obligations, then the probability of his attendance in this scenario would be 90%; however, if the sole reason for his absence is due to anger towards Esra, where he does the opposite of what Esra does, then the probability of his attendance would be 100%, since their current absence from the party would indicate that Ali is not angry. Therefore, the conditional probabilities of observed variables alone are insufficient for answering counterfactual queries accurately; one requires additional information about the mechanisms responsible for these probabilities.

At this point, Balke & Pearl (1994) introduce a functional specification that models the influence of a variable E on a variable A using a deterministic function: $a = F_a(e, \epsilon_a)$. The unknown factors that may affect variable A are represented by the error term, exogenous variable(s), or as they call it, *disturbance* ϵ_a. The prior probability distribution $P(\epsilon_a)$ determines the likelihood of these factors. For instance, ϵ_a may comprise various components such as Ali's anger towards Esra and Ali being grounded by his parents. When a specific value is assigned to ϵ_a, variable A becomes a function of E. Consequently, each value within the domain of ϵ_a can be used to specify a *response function* that maps every value of E to a value in the domain of A. Although the error term ϵ_a may have several components, it can always be replaced by an equivalent variable that is minimal. To achieve this, the domain of the error term is partitioned into equivalence regions, each of which corresponds to a single response function, as outlined by Pearl (1993). These equivalence classes, in a formal sense, can be defined by a function $r_a : \text{dom}(\epsilon_a) \rightarrow \text{N}$. The function r_a assigns a value to each element in the domain of ϵ_a based on the following criteria: i) if $F_a(e_0, \epsilon_a) = 0$ and $F_a(e_1, \epsilon_a) = 0$, then $r_a(\epsilon_a) = 0$; ii) if $F_a(e_0, \epsilon_a) = 0$ and $F_a(e_1, \epsilon_a) = 1$, then $r_a(\epsilon_a) = 1$; iii) if $F_a(e_0, \epsilon_a) = 1$ and $F_a(e_1, \epsilon_a) = 0$, then $r_a(\epsilon_a) = 2$; and if $F_a(e_0, \epsilon_a) = 1$ and $F_a(e_1, \epsilon_a) = 1$, then $r_a(\epsilon_a) = 3$. The function r_a can be seen as a random variable that takes on as many values as the number of functions between variables E and A. Balke & Pearl (1994) refer to this domain-minimal variable as a *response-function variable*, and liken it to the *potential response variable* in Rubin's (1974) model of counterfactuals, which were developed to support causal inference in statistical analysis (Balke & Pearl, 1993; Rubin, 1974 as cited in Balke & Pearl, 1994).

The reason we touched upon this functional specification by Balke & Pearl (1994) was to provide a basic explanation for their approach to representing counterfactuality. They claim that it is insufficient to calculate the posterior distribution of each disturbance variable (ϵ) independently and treat them as unrelated quantities. Even though the disturbance variables are initially independent, the observed evidence tends to create dependencies among the parents of the observed variables, and these dependencies need to be considered in the posterior distribution. One effective way to account for these dependencies is through the causal network's structure itself. Thus, to differentiate between the variables in the real and counterfactual worlds, a separate network will be used for each world. In this approach, evidence can be instantiated on the real-world network, and the solution to the counterfactual query can be determined as the probability of the counterfactual consequent computed in the counterfactual network where the counterfactual antecedent is enforced. The crucial question to ask is, how are the networks for the real and counterfactual worlds linked? Since any exogenous variable, ϵ_{a}, is not influenced by forcing the value of any endogenous variables in the model, the value of that disturbance will be identical in both the real and counterfactual worlds. Therefore, a single variable can represent the disturbance in both worlds. ϵ_a thus becomes a common causal influence of the variables representing A in the real and counterfactual networks, respectively, which allows evidence in the real-world network to propagate to the counterfactual network.

In light of these, let us get back to the party example and say that we have *observed* that Ali is not present at the party, if he *were*, *would* he and Veli *have* fought? Figure 13 is an illustration of the model created following Balke & Pearl's (1994) notation and it displays the revised BN constructed for the party example.

The probabilities in the model indicate Balke & Pearl's comprehension of the characters involved. Specifically, $P(r_a)$ selection represents their belief that Ali generally attends the party only when Esra is present ($r_a = 1$); however, they also believe that Ali may not attend the party for certain reasons (about 7% of the time) such as illness or parental discipline, and this scenario is represented by $r_a = 0$. Furthermore, Ali may attend the party only when Esra is absent (about 3% of the time) due to a vengeful mood, and this exception is denoted by $r_a = 2$. Additionally, $P(r_f)$ portrays their understanding that there is a small probability (5%) that Ali and Veli would not get into a fight regardless of whether they attend the party ($r_f = 0$). There is also a 5% chance [$P(r_f = 9) = 5\%$] of a fight occurring either inside or outside the party, but not if only one of them is present. Figure 14 is what the model looks like after the real world observations (a_0) and the actions (edge removal, a^*_1) are integrated into the model.

If the evidence is propagated accordingly, the probability of Ali and Veli fighting if Ali were at the party [$P(f^*_1 \mid a^*_1, a_0)$] is revealed to be 0.79, which is in line with Beril's statement that Ali and Veli would have fought if Ali were present at the party, given that he was not. In comparison, the probability for the indicative query about Kerem's thoughts [$P(f_1 \mid a_1)$] is 0.11, implying that if we had observed Ali at the party, then it is likely that Ali and Veli would not have fought. This demonstrates the contrast between the ways to find out the answers to counterfactual and indicative queries.

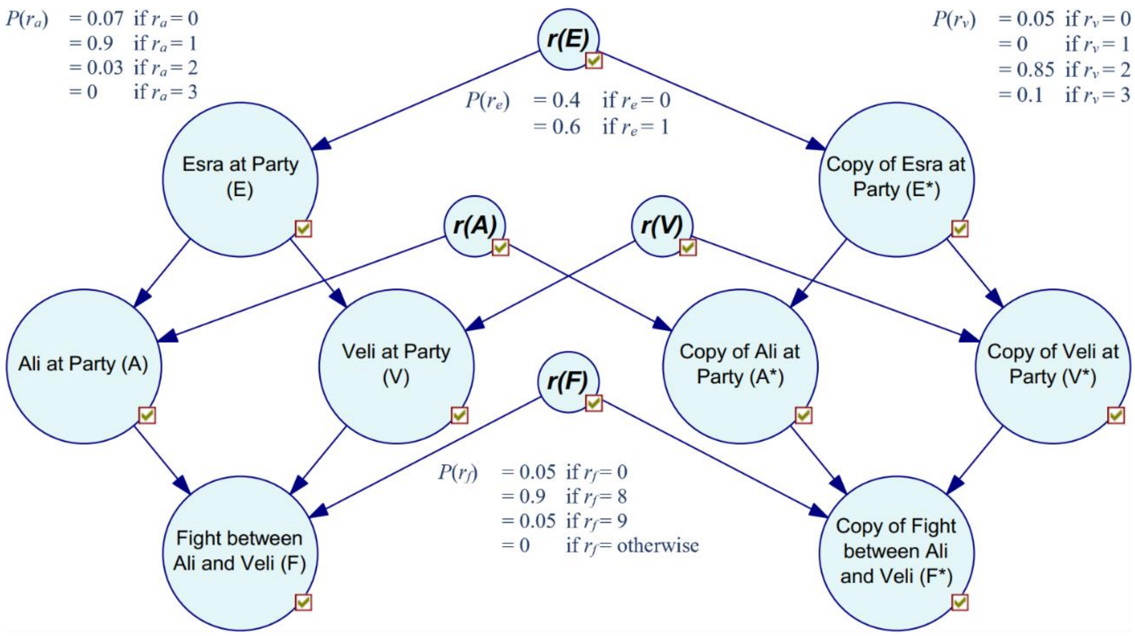

Figure 13: The revised BN that represents the party example. The variables marked with asterisks are the counterfactual counterparts of those that are unmarked.

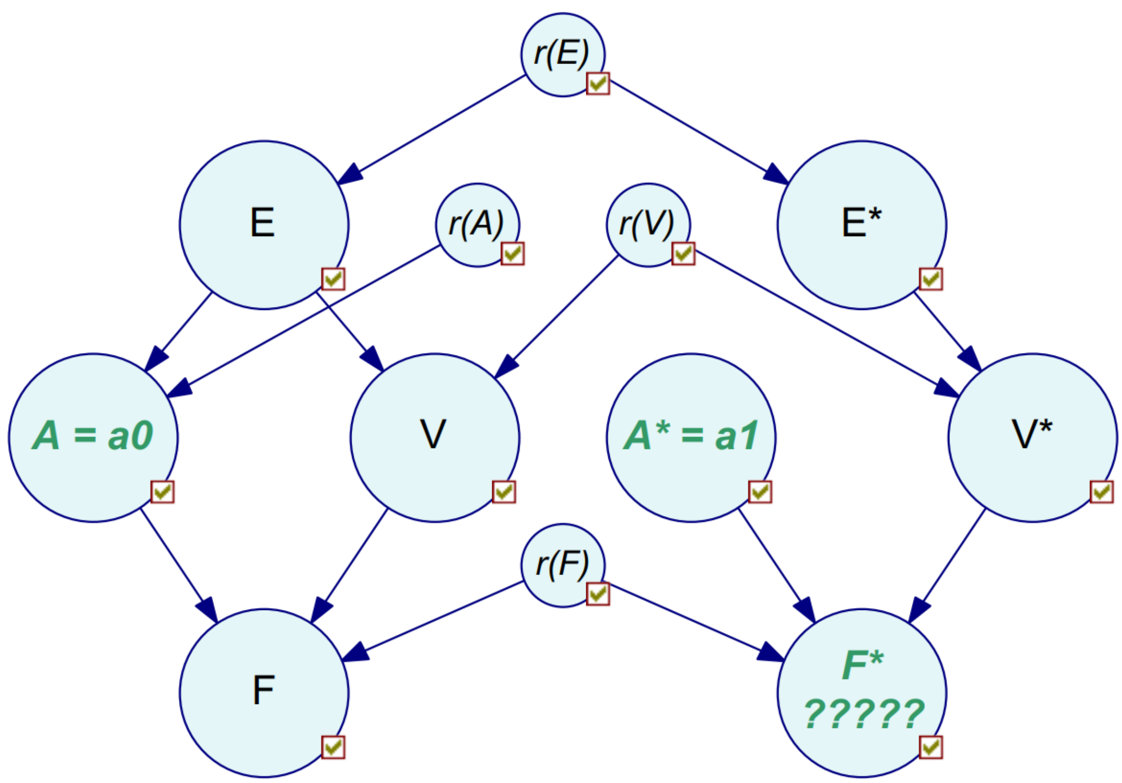

Figure 14: The *re*-revised BN to represent the given scenario: A as a_0, A* as a_1, and all the edges that are directed to A* are removed, and the main focus is on the value of F*.

4.2. Revisiting the "Firing Squad"

Given the aforementioned information, it would be beneficial to revisit the Firing Squad example (3.3.1) and to re-examine the questions that were addressed by considering the implications of *-DIysA* and *-sAydI* in the analysis.

4.2.1. On Rung Correspondence

The illustration in Figure 15 is a composite of the three figures that were utilized in the representation of the Firing example. The diagram on the left represents the default BN, the one in the middle depicts the scenario of intervention, and the diagram on the right illustrates the counterfactual example.

As previously discussed, the first rung of the Ladder of Causation (see 3.1.1) pertains to associations and it is within this realm that inquiries such as *If the prisoner is dead, was the court order issued?* can be addressed. The BN that allows us to answer this question (as yes) is the one on the left in Figure 15, and the Turkish counterpart to the question is in (72); notice how the subordinate verb is marked with *-DIysA*.

(72) [Esir öl**düyse]** mahkeme kararı çıkmış mıdır?

[If the prisoner is dead (has died/has been killed)], was the court order issued?

The second rung (see 3.1.2) is concerned with interventions and it is in this capacity that questions such as *What happens to the prisoner if Soldier A decides on their own initiative to fire (without waiting for the captain's command)?* can be investigated. The BN that allows us to answer this question (as *they will be dead*) is the one in the middle in Figure 15, and the Turkish counterpart to the question is in (73); notice how the subordinate verb is marked with the aorist + *-(y)sA*.

(73) [Asker A (komutandan emir gelmeden) kendi kendine ateş etmeye karar ver**irse**] esire ne olur?

What happens to the prisoner [if Soldier A decides on their own initiative to fire (without waiting for the captain's command)]?

The third rung (see 3.1.3) of the Ladder of Causation pertains to counterfactual statements and it is within this domain that inquiries such as *What would have happened to the prisoner if Soldier A had decided not to fire?* can be examined. The BN that allows us to answer this question (as *they would be dead*) is the one on the right in Figure 15, and the Turkish counterpart to the question is in (74); notice how the subordinate verb is marked with *-sAydI*.

(74) [Asker A ateş etmemeye karar ver**seydi**] esire ne olurdu?

What would have happened to the prisoner [if Soldier A had decided not to fire]?

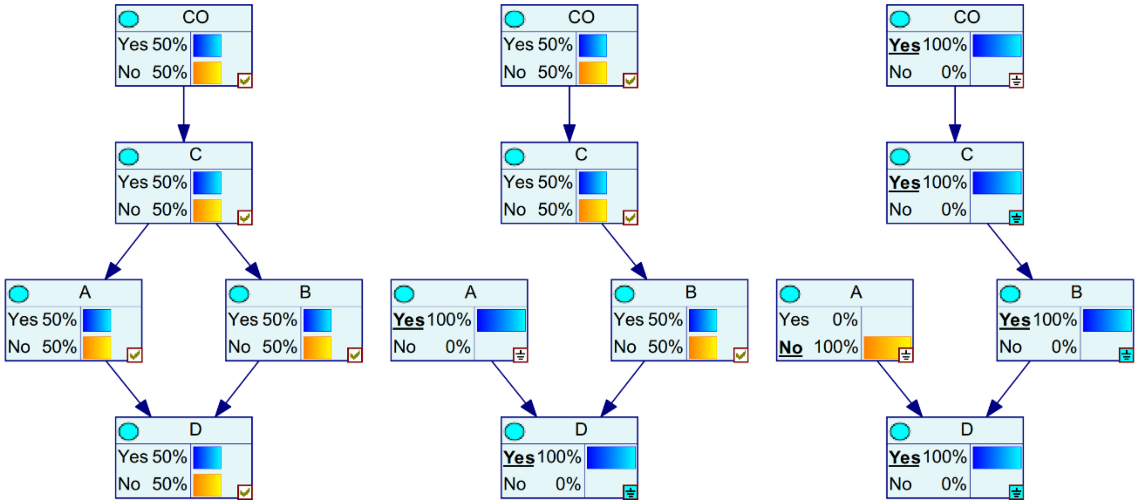

Figure 15: The Firing Squad figures combined. From left to right: the default BN, intervention example, counterfactual example.

As a preliminary conclusion, we can say that the distinction between *-DIysA* and *-sAydI* is not limited to syntax and semantics, it is event-referential, in that the former aligns with level 1 in Pearl's framework, while the latter corresponds to level 3, and also representational, in that the event-referential difference necessitates these constructions to be represented through two distinct BNs.

4.2.2. Computing Counterfactuals

The aim of discussion on the Firing Squad in 3.3.1 was to introduce the representational notation. We discussed another example above, the Party example from Balke & Pearl (1994), to show what and how calculations can be made regarding counterfactuality. Now, let us look at the Firing Squad example taking what we discussed in the Party example into account, and while we are at it, let us assign different probabilities to the variables.

Figure 16 is an illustration of the Firing Squad example with the new CPTs.

Figure 17 represents the observation example with the new CPTs, and lets us answer question (72), the probability of the court order being issued given that the prisoner is dead, as 0.34.

Figure 18 represents the intervention example with the new CPTs, and lets us answer question (73), the prisoner's state given that Soldier A decided on their own initiative to fire, as *they would be 92% dead*.

Now, we could represent the counterfactual example with the new CPTs through a BN which looks like the one in Figure 10. That BN was an adequate and intuitively easy to understand one for representation; however, it does not hold when one wants to do certain calculations within the model. We will be amalgamating what we have shown here and what we have introduced in the party example to be able to do just that.

Taking Balke & Pearl's (1994) approach into account and adding what Pearl (2009) suggested later on as *The Twin Network Method*, we can arrive at Figure 19. We use identical networks, one for the actual world and another for the hypothetical one, that are connected through the same exogenous variables, which in this case is just the Court Order (CO), and we add asterisks (*) to the labels of the variables in the hypothetical world to differentiate them. Let us assume that we observe that the prisoner is dead, and we know that the court order was not issued, what would have happened if A had not shot? The prisoner would have been 93% alive.

The way to represent the potential effect of another exogenous variable, the Nervousness of Soldier A, can be seen in Figure 20: N is added as a common parent for A and A*.

Now, let us imagine a scenario in which we observe that the prisoner is dead, and we wonder what would have happened to them if a highly nervous (N = high) versus a calm (N = low) A had not shot. Comparing the probabilities for CO and D* in Figure 21 and Figure 22, we can see that observing a highly nervous A when the prisoner is dead and if A had not shot means that the probability of CO being issued is 56% and the probability of the prisoner's death is 52%, whereas observing a calm A in the same situation means that the former probability is 59% and the latter is 54%.

We have explored another way to represent counterfactuality, and it is time to turn our attention back to Turkish with another example in whose exploration we will employ this twin network approach.

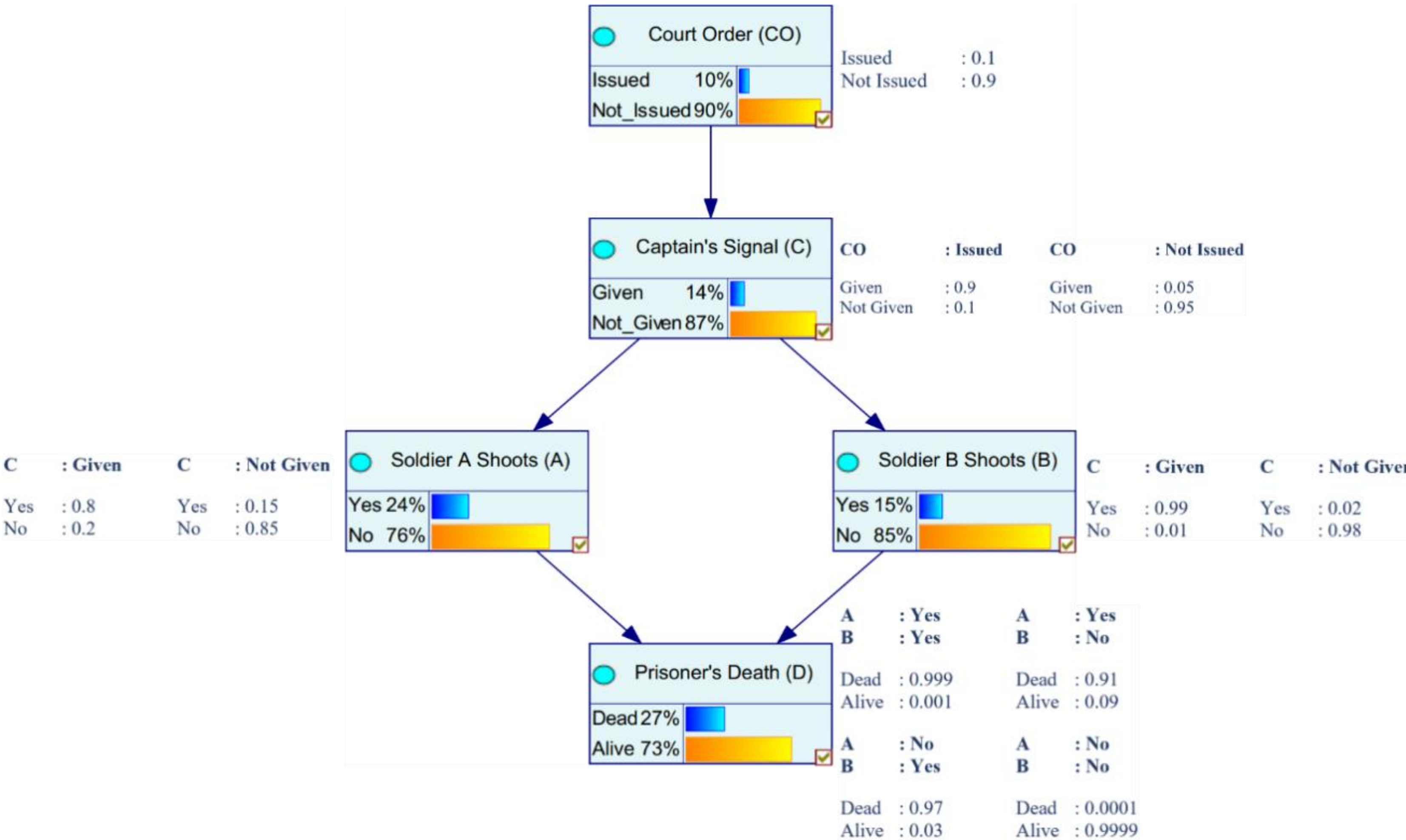

Figure 16: The Firing Squad with new CPTs.

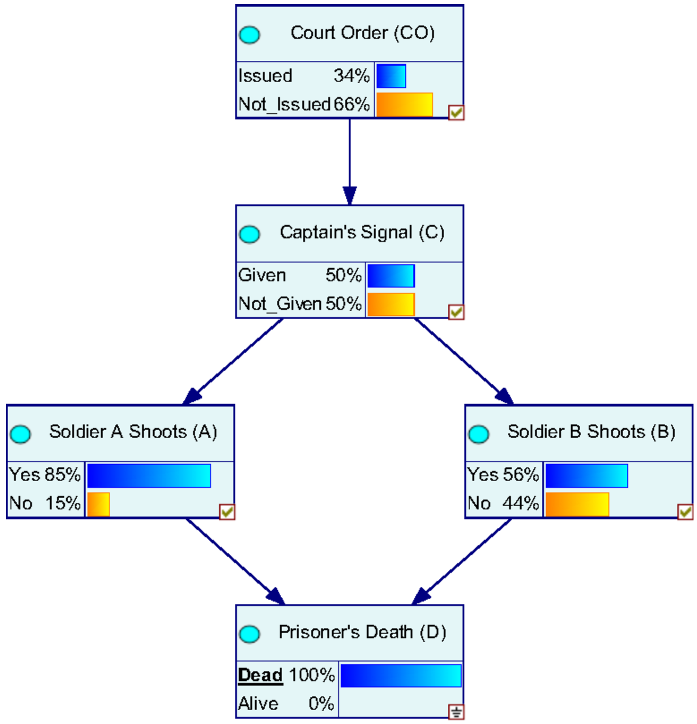

Figure 17: The observation example with new CPTs.

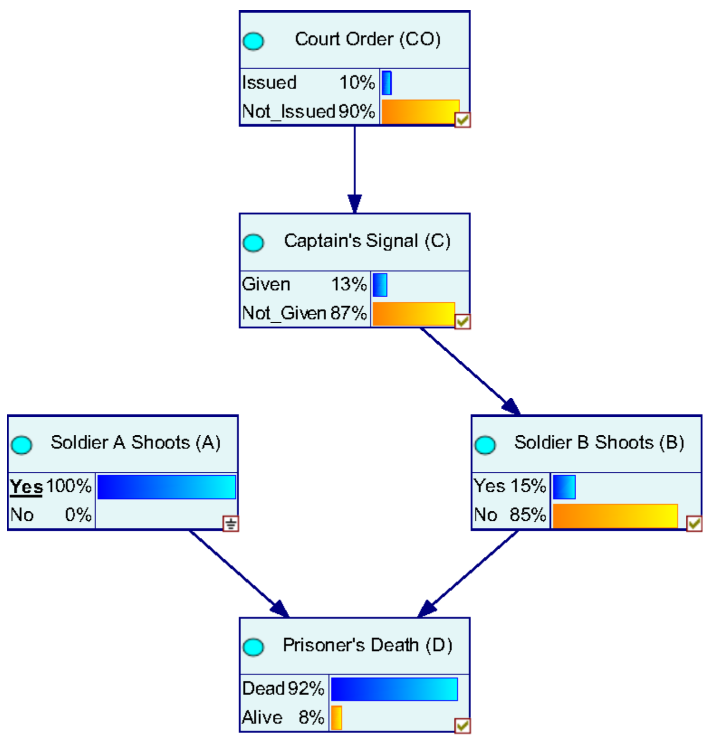

Figure 18: The intervention example with new CPTs.

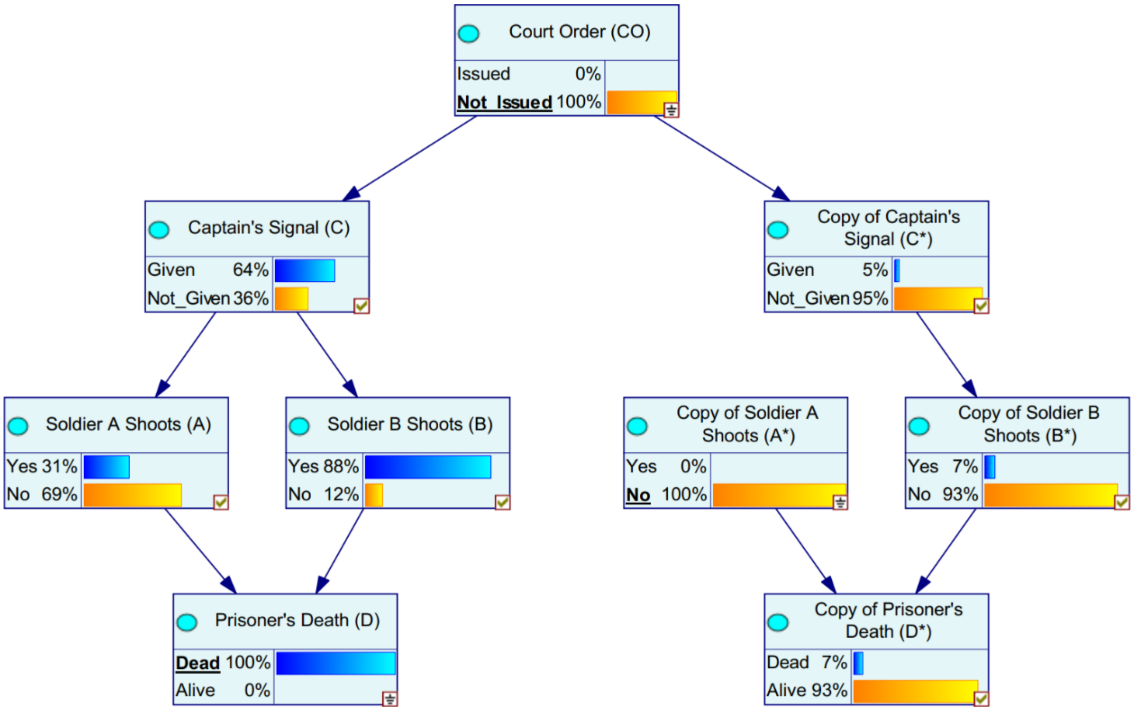

Figure 19: The twin network approach to the Firing Squad example.

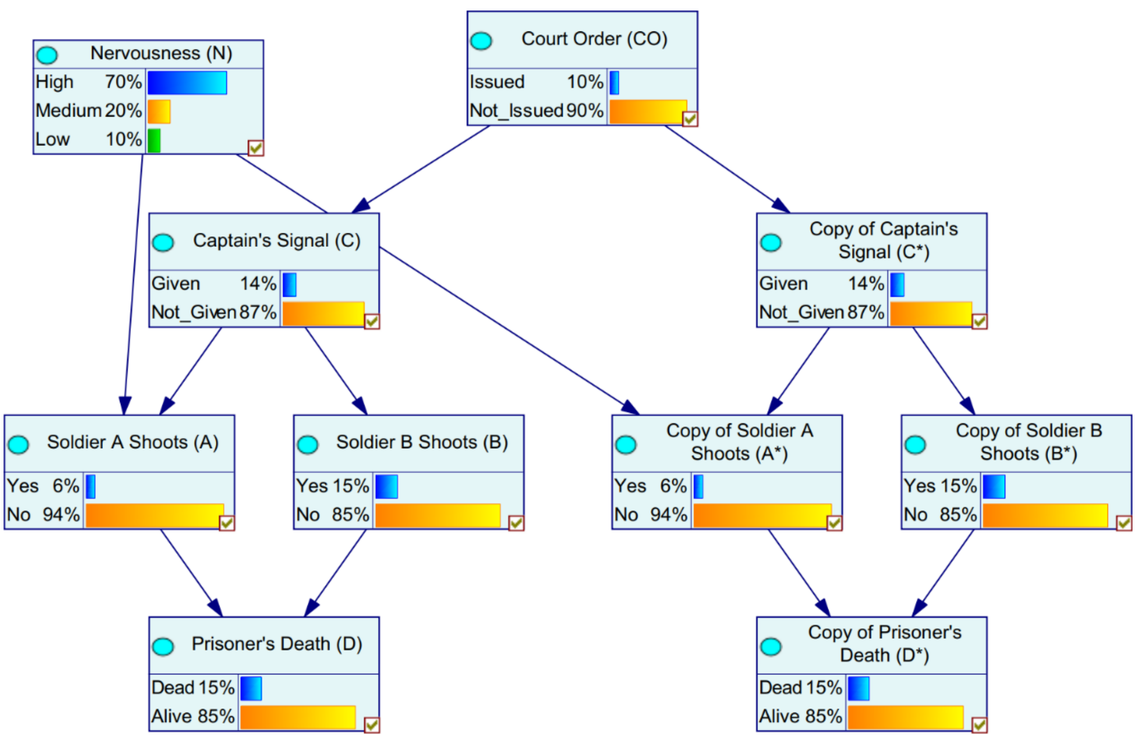

Figure 20: Adding another exogenous variable to the Firing Squad: the Nervousness (N) of Soldier A (A).

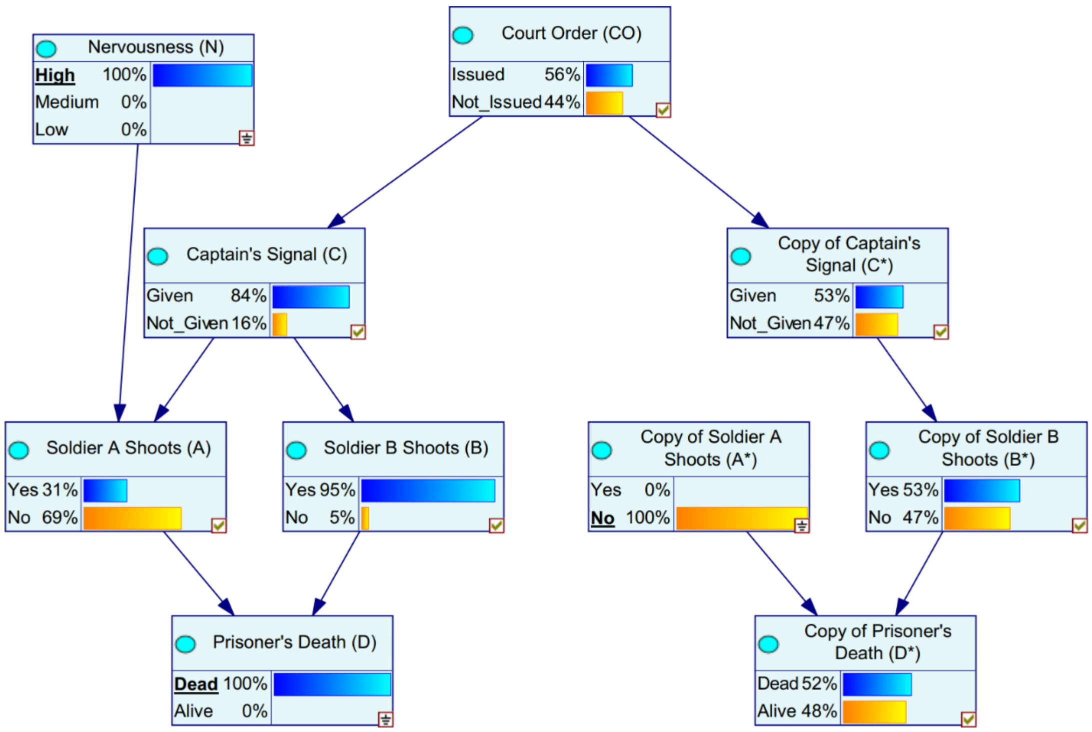

Figure 21: Observing a highly nervous A and a dead prisoner. A had not shot in the hypothetical world, we are looking at the probabilities for CO and D*.

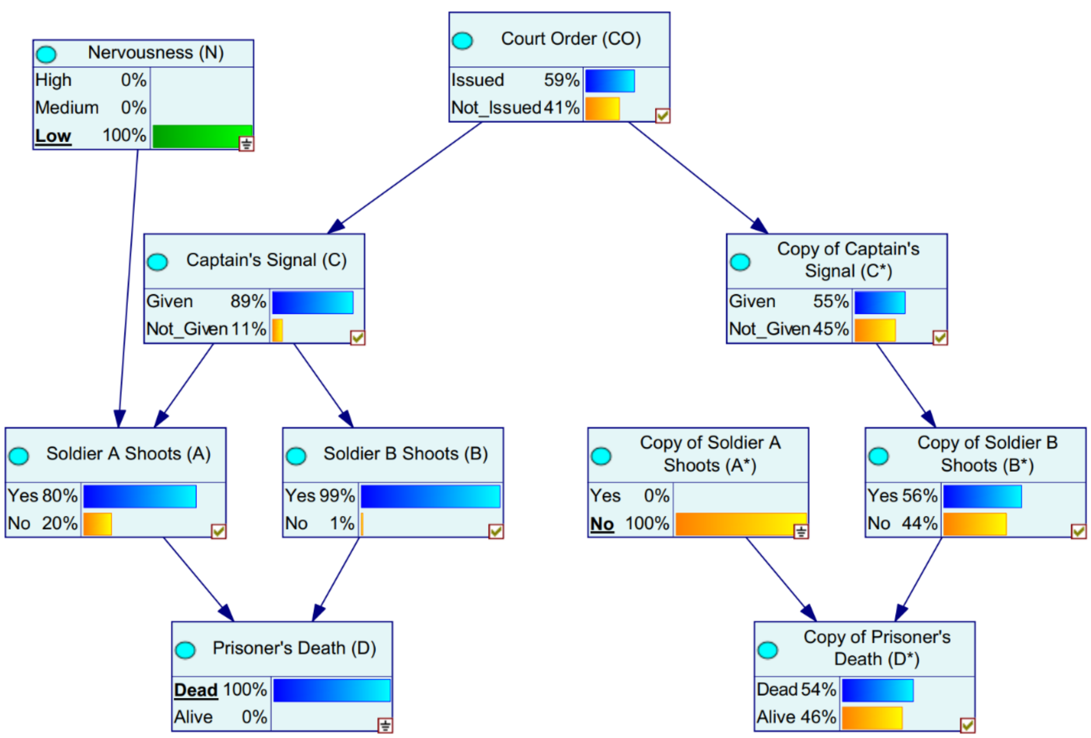

Figure 22: Observing a calm A and a dead prisoner. A had not shot in the hypothetical world, we are looking at the probabilities for CO and D*.

4.3. “Finding the Medicine”

According to Ruhi et al.[15](2000), the example sentence in (75), displays a counterfactual conditional and (76) a hypothetical conditional. As per the authors, (77a) and (77b) constitute infelicitous[16] continuations to (75) and (77c) constitutes an infelicitous continuation to (76).

(75) [İlacı bulma**saydı**] hastalığı ilerleyecekti.

[If they had not found the medicine], their disease would have worsened.

(76) [İlacı bulma**dıysa**] hastalığı ilerlemiştir.

[If they did not find the medicine], their disease must have worsened.

(77)

a. Infelicitous continuation to (75)

Umarım ilacı bulmuştur.

Hope *they found the medicine.*

b. Infelicitous continuation to (75)

Belki de ilacı bulmuştur.

Maybe *they found the medicine.*

c. Infelicitous continuation to (76)

İyi ki ilacı buldu.

Luckily *they found the medicine.*

In hypothetical conditional sentences, the antecedent does not possess a discernible truth value (Oktar & Can, 2016). Consequently, as the proposition *umarım ilacı bulmuştur* (hope they found the medicine) in (3a) expresses a wish and a possibility, the lack of an opposite truth value within the antecedent does not result in a contradiction or infelicity. Conversely, counterfactual statements presuppose that the proposition within the antecedent clause is false (e.g., Anderson, 1951; Barwise, 1986; Pearl et al., 2016, among many others); therefore, the negative suffix *-mA* in the antecedent of (77c) denotes that

[15] Note that the original work is in Turkish and the exact terms they use are *gerçek karşıtı* (unreal) and *varsayımsal* (hypothetical). See the original works footnotes for their explanation.

[16] Ruhi et al. (2000) call them *kabul edilemez* (unacceptable), and Can Bakırlı (2010) calls them *dilbilgisi dışı* (ungrammatical); we opted for the more general *infelicitous*.

the opposite of the proposition holds true, specifically, that the action of *bul-* (find) is factual and thus, the medicine was found.

Be that as it may, consider a scenario in which person A and person B are talking about a certain person C, a mutual friend who, as far as A and B are concerned, was very sick and searching for a hard-to-find medication at the time of their last encounter. Now, suppose A and B are not the best friends one could ask for and they have been out of touch with C for an extended period. Providing such a context for the sentences in (75) renders u*marım ilacı bulmuştur* (hope they found the medicine) in (77a) and *belki ilacı bulmuştur* (maybe they found the medicine) in (77b) much more plausible than they are without a context. A and B are cognizant of the fact that *C's disease would have worsened if they had not found the medicine*; but then again, they do not possess knowledge of whether or not C ultimately found it or not. That adding *bildiğim kadarıyla* (*as far as I know/to the best of my knowledge*) sentence-initially as in (78) or *ne yaptı acaba* (*I wonder what they did*) sentence-finally as in (79) or both as in (80) does not cause any infelicity proves this point. It is highly probable that these additions are acceptable because they introduce a particular *uncertainty* as to whether the drug was found or not, which can be said to be another way of entering *soft evidence,* a concept we will explain in the upcoming subchapter.

(78) **Bildiğim kadarıyla**, [ilacı bulmasaydı], hastalığı ilerleyecekti.

__As far as I know/To the best of my knowledge__, [if they had not found the medicine], their disease would have worsened.

(79) [İlacı bulmasaydı], hastalığı ilerleyecekti. **Ne yaptı acaba?**

[If they had not found the medicine], their disease would have worsened. ***I wonder what they did (i.e., if they found it or not).***

(80) **Bildiğim kadarıyla,** [ilacı bulmasaydı], hastalığı ilerleyecekti. **Ne yaptı acaba?**

__As far as I know/To the best of my knowledge__, [if they had not found the medicine], their disease would have worsened. ***I wonder what they did (i.e., if they found it or not).***

Consider also a later conversation between A and C, who seems to have recovered, as in (81); such a conversation should not be possible according to the accounts that suggest that the proposition within the antecedent clause is presupposed to be false, yet it is.

(81) A: İlacı bulmasaydın hastalığın ilerleyecekti; nasıl iyileştin, ilacı buldun sanırım?
Your disease would have worsened/was going to worsen if you hadn't found/didn't find the medicine; how did you recover, I suppose you found the medicine?

C: İlacı bulamadım ama çok iyi bir cerrah tarafından ameliyat edilme fırsatım oldu.
I could not find the medicine, but I had the opportunity to be operated on by a very good surgeon.

Another curious aspect of the first parts of the sentences in (75) and (76) is the way they interact with negation. In this case, removing the negation from the antecedent's verb seems to require the consequent's verb either to be negated, as in (82), or to be changed accordingly, as in (83), to give a similar meaning; this might be related to *conditional perfection* (strengthening of a conditional to biconditional).

(82) a. [İlacı bulsaydı] hastalığı ilerlemeyecekti.
[If they had found the medicine], their disease would not have worsened.
b. [İlacı bulduysa] hastalığı ilerlememiştir.
[If they found the medicine], their disease must not have worsened.

(83) a. [İlacı bulsaydı] iyileşecekti.
[If they had found the medicine], they would have recovered.
b. [İlacı bulduysa] iyileşmiştir.
[If they found the medicine], they must have recovered.

4.4. Modeling Counterfactual Statements in Turkish: The Proposed Approach

4.4.1. On Contextualization

Since Finding the Medicine example above is *context-free*, the BN that corresponds to it would only be a two-node one, as in the diagram on the left in Figure 23, which would be, to say the least, *uninformative*. We could still represent a potential counterfactuality as we did for the party example, displayed through the diagram on the right in Figure 23, but it would not really be the most ideal model for the reasoning scheme we follow.

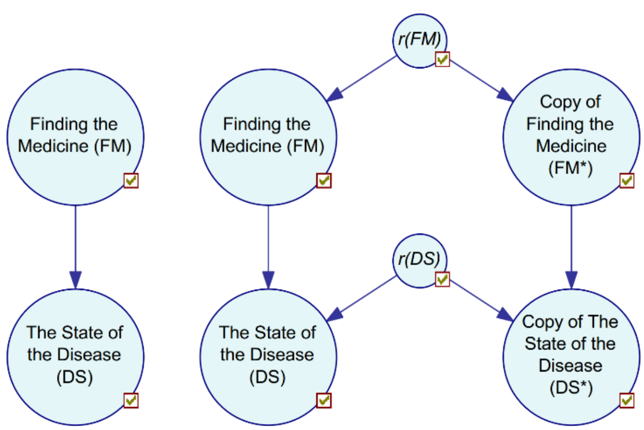

Figure 23: On the left, the potential, uninformative BN that represents Finding the Medicine example; on the right, the same BN with exogenous variables and counterfactual counterparts in a manner similar to Balke & Pearl (1994).

As they would most likely consist of a single-event antecedent and a single-event consequent, any given out-of-context conditional/counterfactual sentence would be as delimiting and uninformative as the Finding the Medicine example in terms of representability, which brings us to the first step of our approach. We believe there is merit in enriching the context to be able to get a clearer picture of the relationship between the antecedent and the consequent. The variables for an example of the said enrichment are defined below, and the resulting diagram is demonstrated in Figure 24.

- *Economic Crisis*: The probability of a global economic crisis; *yes* or *no* ~ exists or does not exist.
- *Finding the Medicine*: The probability of finding the medicine; *found* or *not found*. It is meant to be a hard-to-find medicine but there is of course more than one way to find it. Immediately depends on *Economic Crisis*.
- *Financial Health*: The financial wellbeing of a potential medicine buyer; *wealthy*, *average*, or *poor*. Immediately depends on *Economic Crisis*.
- *Obtaining the Medicine*: The probability of obtaining the medicine; *obtained* or *not obtained*. It is meant to refer to a potential buyer completing the purchase in question but that is not the only way to obtain it. Immediately depends on *Finding the Medicine* and *Financial Health*.
- *The State of the Disease*: The state of the disease going forward; *better* or *worse*. Immediately depends on *Obtaining the Medicine*.

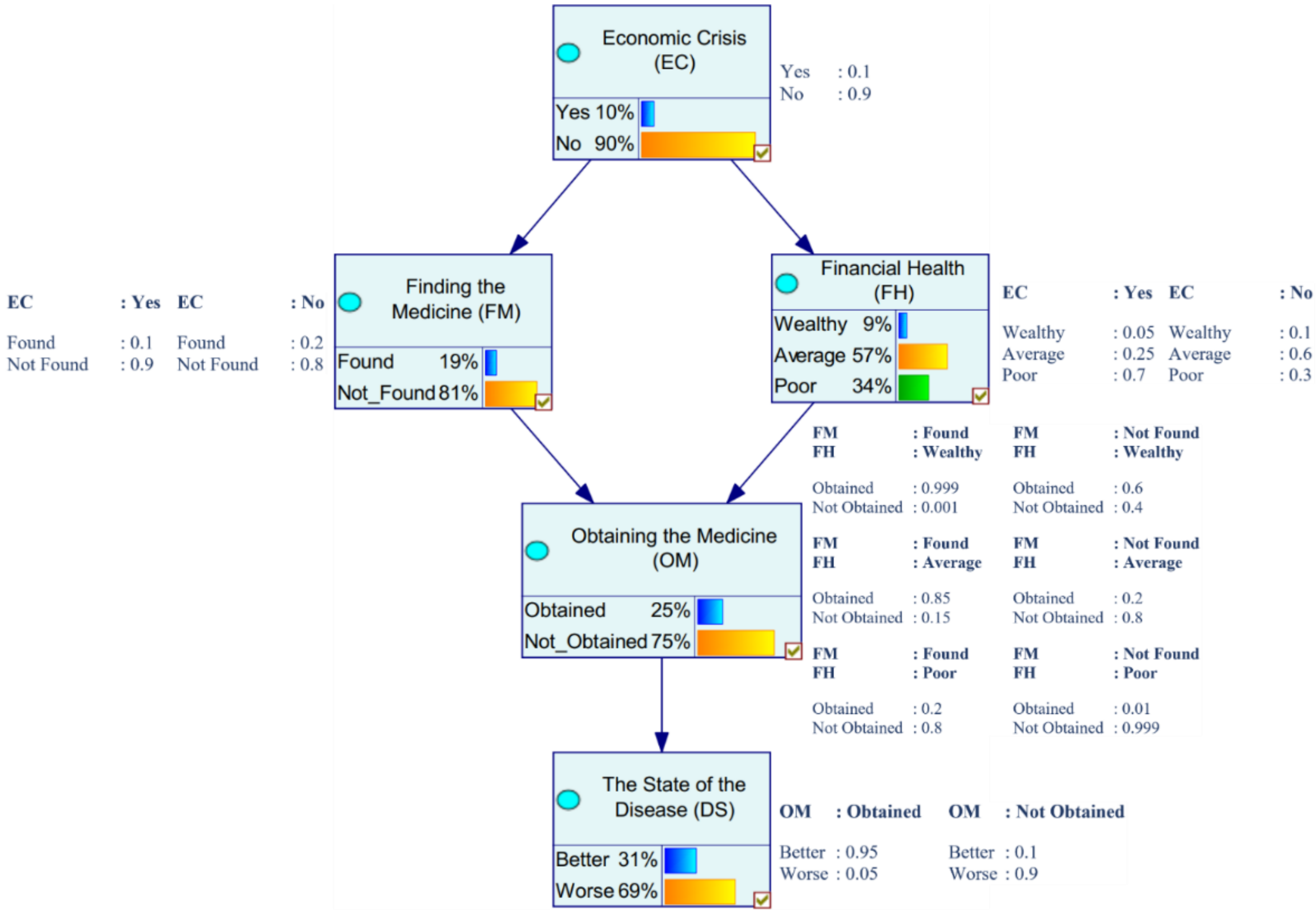

Figure 24: The BN that illustrates an enriched version of Finding the Medicine example.

4.4.2. On Representation

Now, to represent the observation example with *-DIysA* in (76), which corresponds to rung 1 in Pearl's Ladder of Causation, we would simply mark FM as *Not Found* as in the diagram on the left in Figure 25. FM is marked as *Found* in the diagram on the right in Figure 25, which represents the example sentence (82b) or (83b) depending on how negation is interpreted. The difference between the two sentences, and in turn, two diagrams is apparent when DS probabilities are compared.

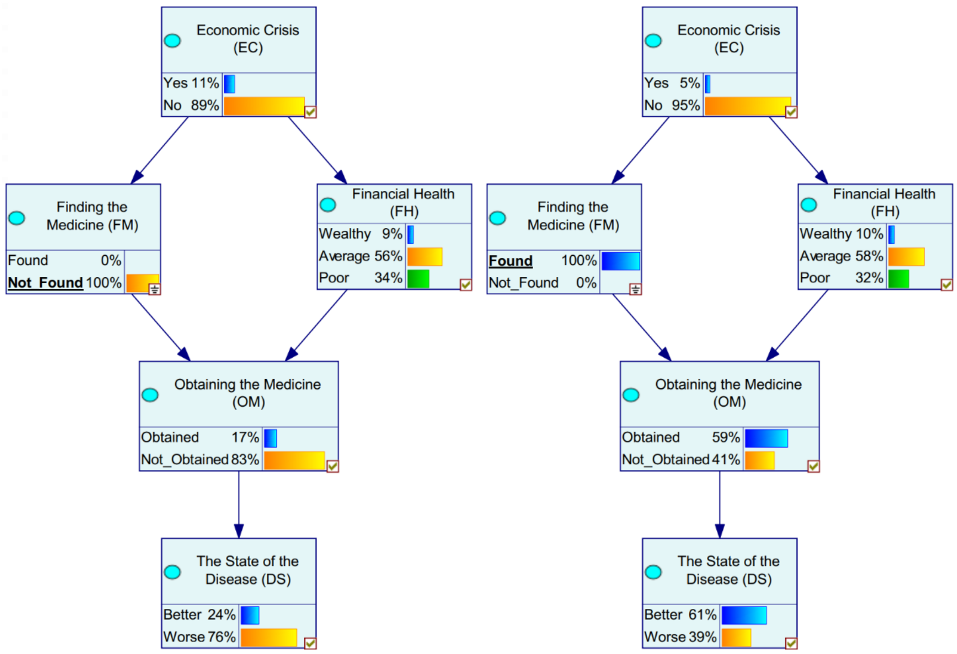

Figure 25: The BN that depicts the observation example with *-DIysA*, (76), on the left; the BN that represents (82b) / (83b) on the right. Compare the DS probabilities.

(84) below exemplifies an intervention with *-(A/I)r* + *-(y)sA*, corresponding to rung 2 in the Ladder, and it can be illustrated as in Figure 26.

(84) [İlacı bulur**sa**] hastalığı ilerlemez (iyileşir?).
[If they find the medicine], their disease would not worsen (they would recover?).

Displaying the counterfactuality in (75) is not a clear-cut task at all. Let us start by assuming that the antecedent in a counterfactual statement needs to be false or unrealized: the medicine is found, obtained, and the disease got better; now, what if it was *not* found? Figure 27 illustrates this scenario.

Figure 28 depicts another scenario in which no observations are made and the hypothetical scenario in which the medicine is not found is entertained. Compare the DS* probabilities in Figure 27 and Figure 28.

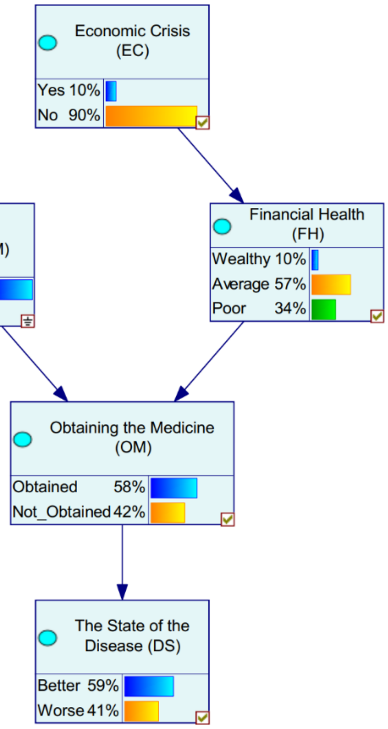

Figure 26: The BN that exhibits the intervention example with -(A/I)r + -(y)sA: the edge from EC to FM removed, FM marked as found. See the DS probabilities.

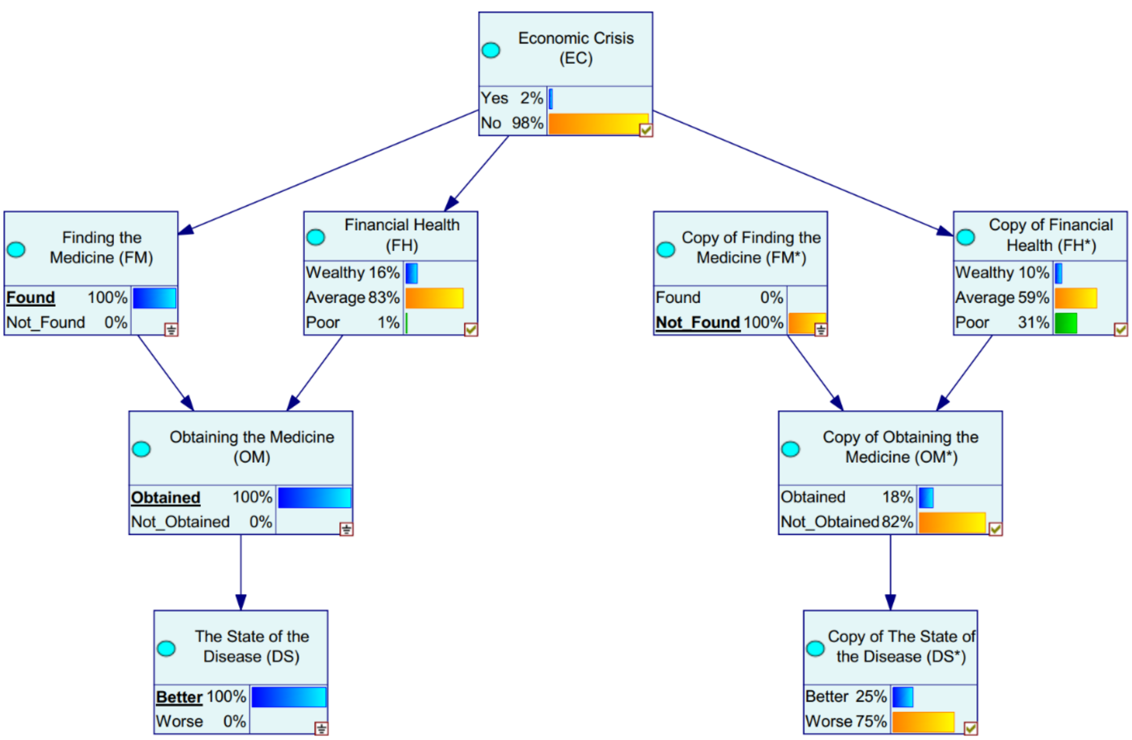

Figure 27: The medicine is found, obtained, and the disease got better; what if the medicine was not found?

The lack of a significant difference between the DS* probabilities may lead one to downplay the value of observations; however, the main idea here is that be it from a grammatical perspective or a graphical perspective, one can, and in fact one frequently *does* in real life, utter sentences such as (75) without any knowledge of the state of affairs regarding the antecedent, which is, in this case, the medicine being found.

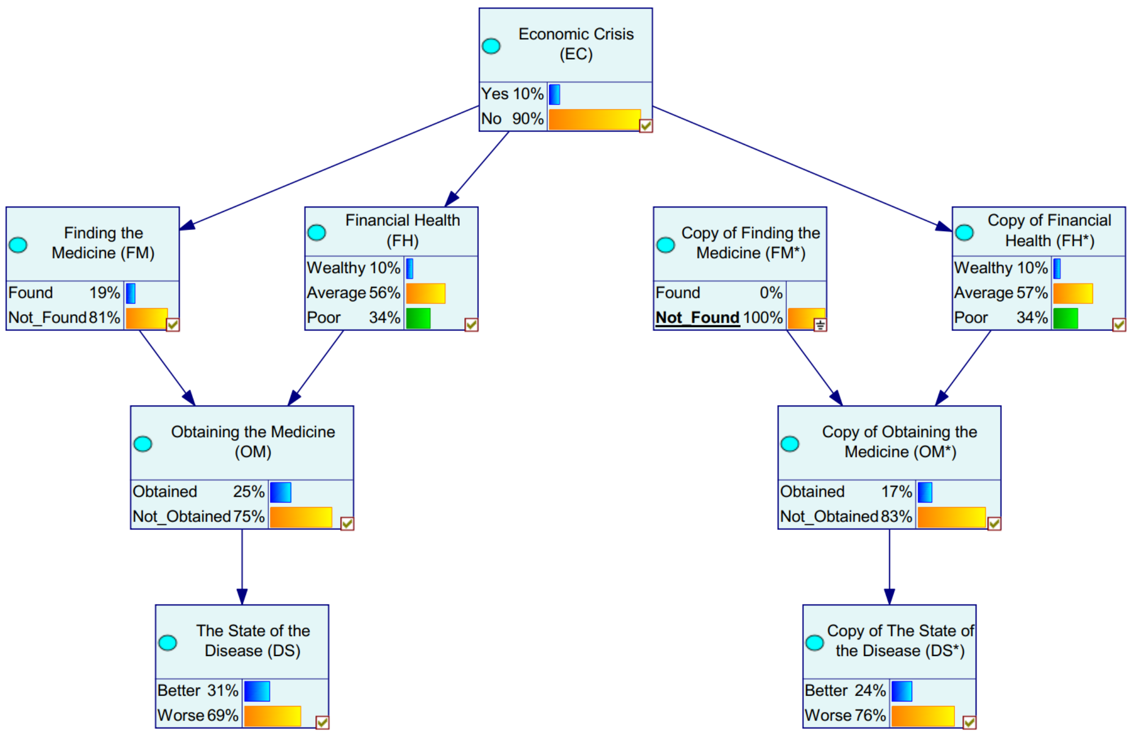

Figure 28: No observations are made, the hypothetical scenario in which the medicine is not found is shown.

4.4.3. On Uncertainty and Exogenous Factors

Up to this point, we made inferences based on *hard evidence*, which refers to the *observation* of a particular event X. By calculating the conditional probability $P(\text{Y} \mid \text{X})$, we can determine the likelihood of another event Y. For instance, in the case of Bernoulli trials, where a coin is flipped multiple times, each flip of the coin is considered hard evidence, as the outcome is definite and can be recorded as either X = 0 or X = 1. Subsequently, traditional Bayesian methods can be used to estimate the probability of the next flip resulting in either heads or tails (Yu, 2021).

In contrast to hard evidence, *soft evidence*[17] is characterized by another probability distribution $R(X)$, as explicated by Peng et al. (2010), or as an *unreliable testimony* of X with a strength of belief in the interval [0, 1], as proposed by Chan & Darwiche (2005) and Jacobs (2019). For example, in the context of Bernoulli trials, if poor visibility conditions make it difficult to definitively ascertain the outcome of a coin flip, we may estimate that there is an 70% probability that it landed heads, i.e., $R(B = 1) = 0.7$. Soft evidence is commonly encountered in various real-world scenarios, such as unreliable witness testimonies in criminal trials (Fields, 2013) and decision-making processes in game theory (Dietrich et al., 2016) (Peng et al., 2010; Chan & Darwiche, 2005; Jacobs, 2019; Fields, 2013; and Dietrich et al., 2016 as cited in Yu, 2021).

Earlier (see 4.3), we introduced some uncertainty as to whether the drug was found or not through adding *bildiğim kadarıyla* (*as far as I know/to the best of my knowledge*) sentence-initially as in (78) or *ne yaptı acaba* (*I wonder what they did*) sentence-finally as in (79) or both as in (80), and likened this modification(s) to entering soft evidence. We can represent this on a diagram through adding a child node[18] as in Figure 29.

Now, compare the diagram on Figure 29 with the diagrams on Figure 30 which show the scenario in which the speaker *knows* and *does not know* if the medicine was found or not. Comparing also the diagrams in Figure 30 with each other, we can see that this way of adding the uncertainty to the network has a significant impact on the DS values.

We mentioned another scenario, (81), in which we have seen an opportunity to be operated on by a very good surgeon, a sort of divine *intervention* if you will. The diagrams corresponding to this scenario would look like the one in Figure 31, in which we added an exogenous variable, Good Fortune (GF), as a parent of DS to represent the unexpected nature of the operation situation.

Now, the patient, despite having failed to find the medicine, looked healthier in (81) because they were lucky to seize that aforementioned opportunity to be operated on by a skilled surgeon. Figure 32 is an illustration of this situation; compare the DS values in Figure 31 and Figure 32.

---

[17] We will be calling the type of uncertain evidence we deal with here *soft evidence*; see Mrad et al. (2015) for a detailed explanation and comparison of different types of uncertain evidence.

[18] Another way to enter uncertain evidence in GeNIe is to use its built-in feature: Set Evidence → Virtual Evidence.

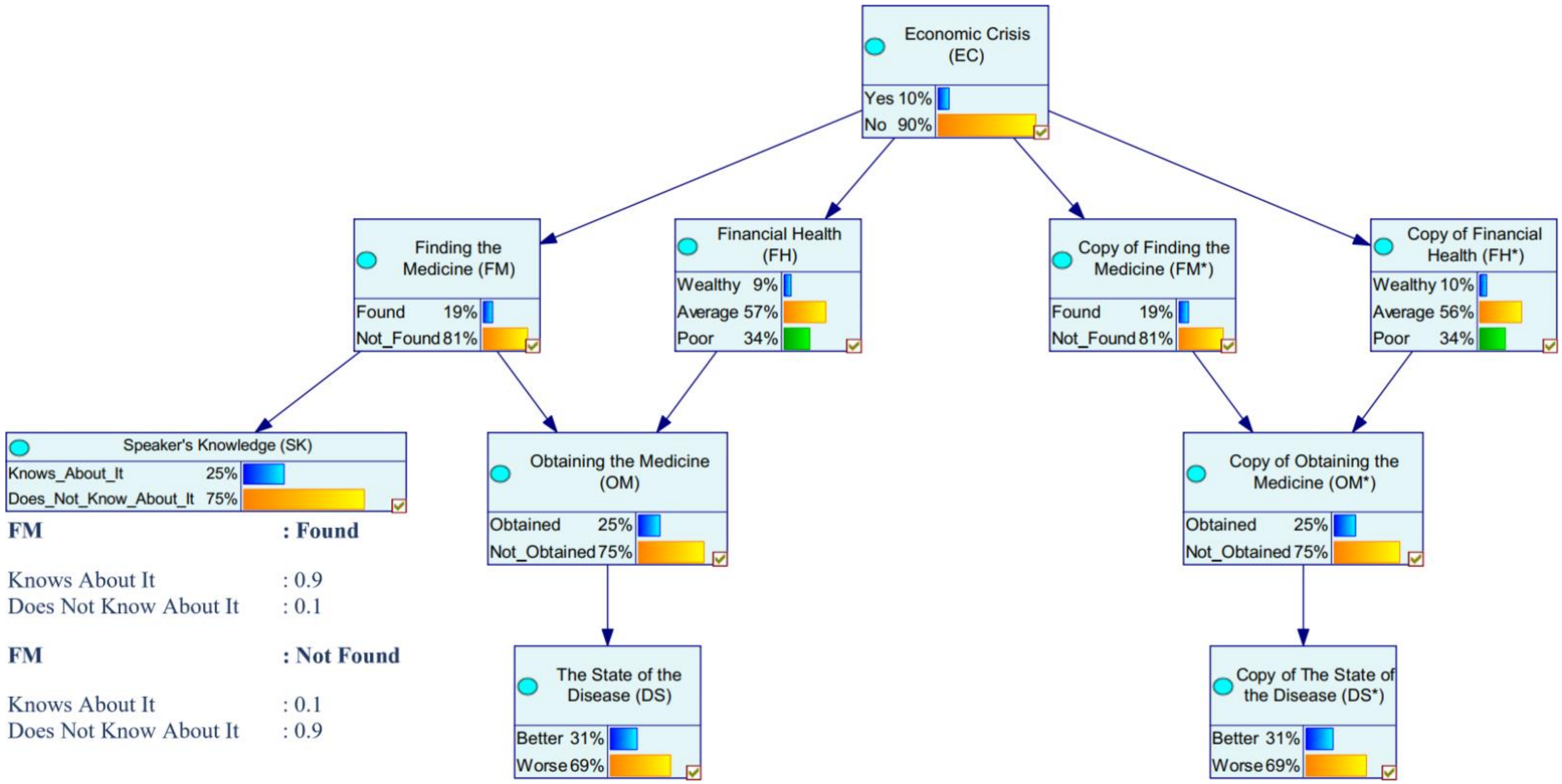

Figure 29: Entering the speaker's knowledge (or lack thereof) as soft evidence, which corresponds to adding it as a child node.

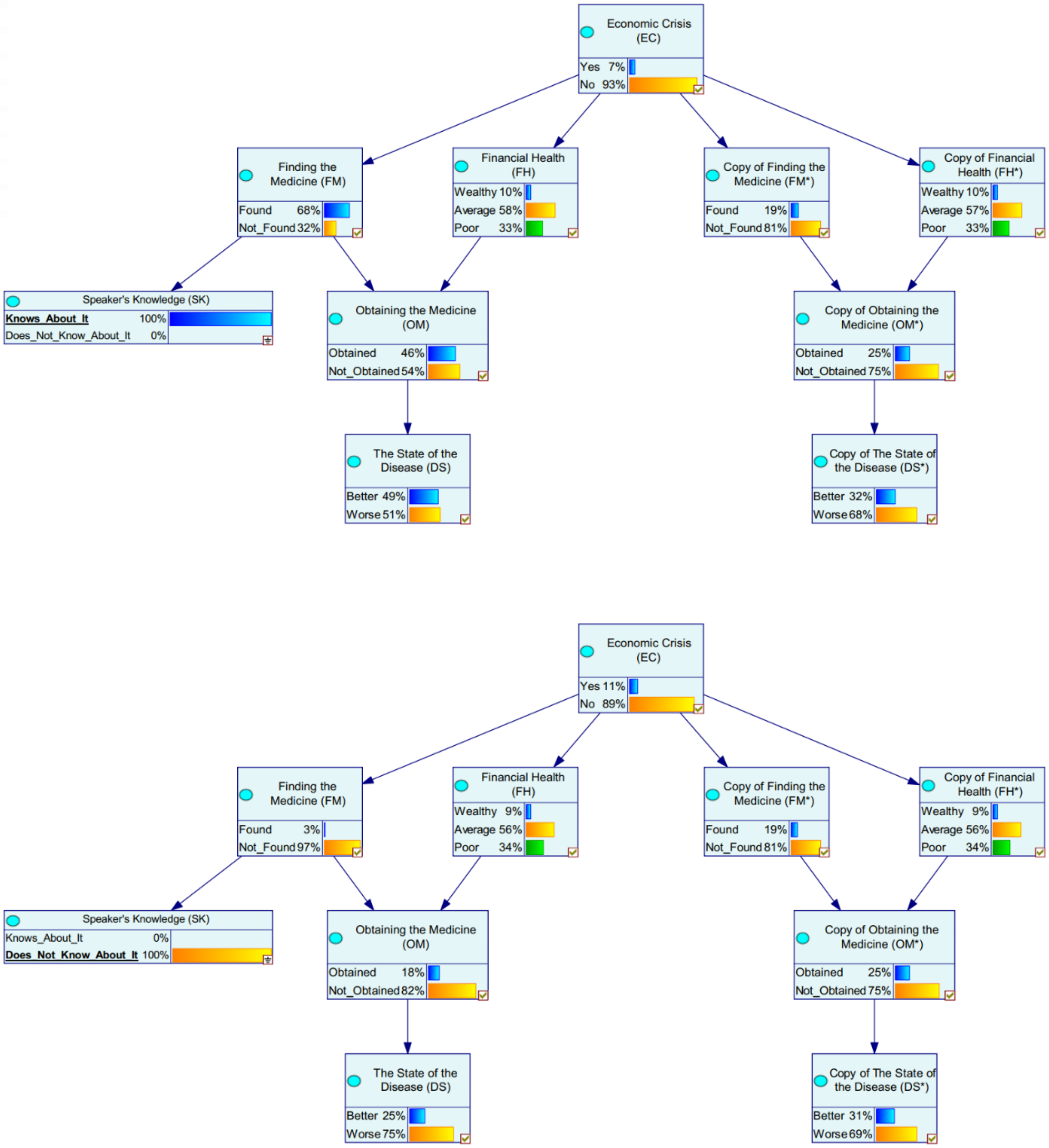

Figure 30: The speaker knows about the medicine situation on the top diagram; and does not know about it on the bottom one. Compare the DS values.

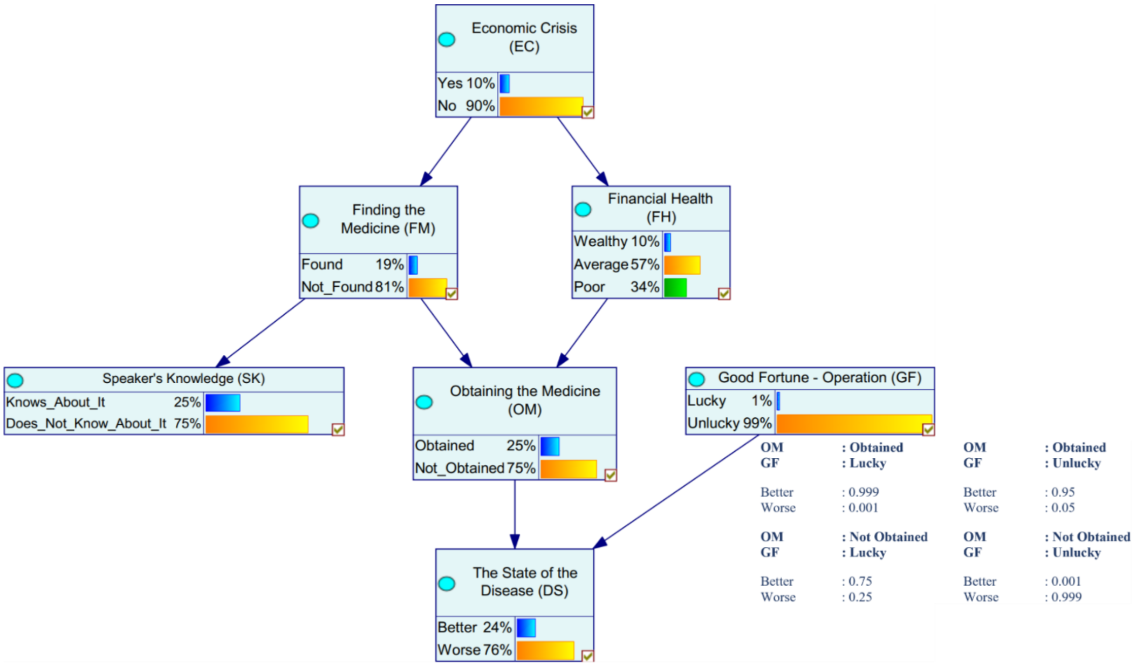

Figure 31: The opportunity to be operated on by a very good surgeon represented through GF.

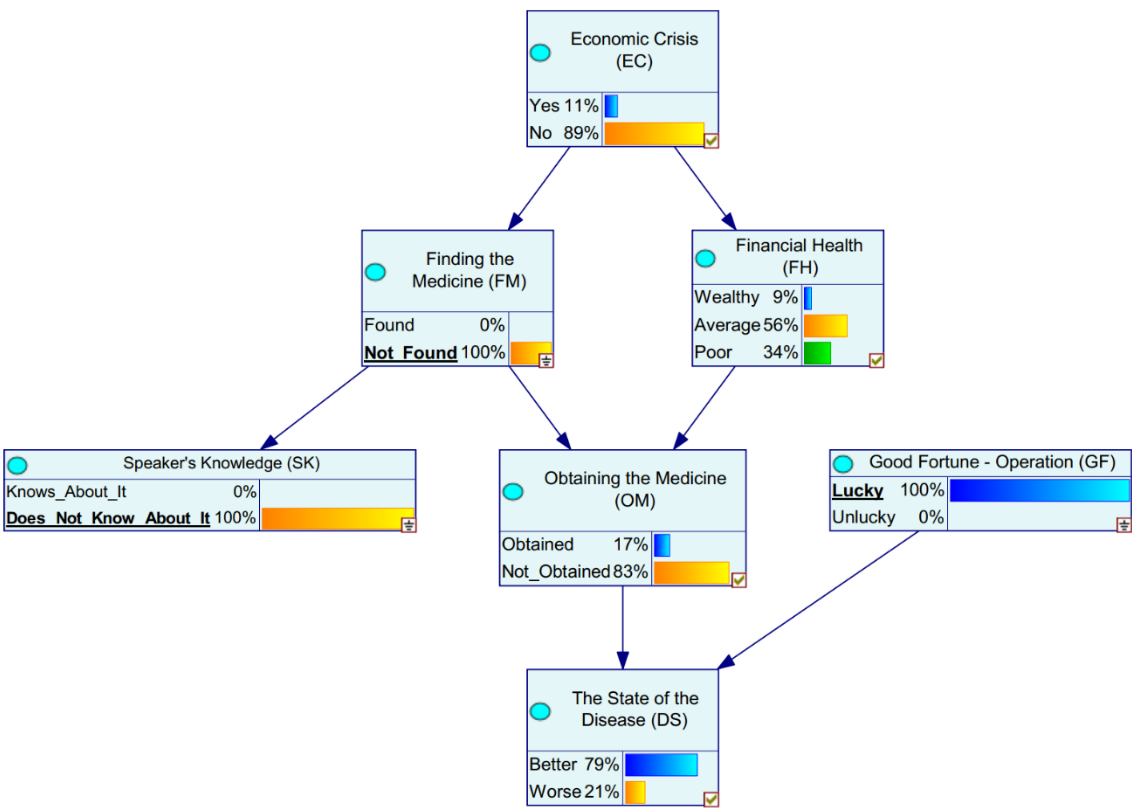

Figure 32: The speaker does not know about the medicine situation and the medicine was, in fact, not found; however, the patient was lucky to find a skilled surgeon.

4.5. Results

The approach we described allows us to represent a plethora of observation, intervention, and counterfactual scenarios, and scenarios with uncertain evidence and/or unexpected, exogenous factors. To summarize, in analyzing a given counterfactual sentence, i) we enrich its context beyond the state of affairs described in the antecedent-consequent duo, ii) define the causal relationships between the products of the said enrichment, iii) assign arbitrary probabilities to them based on what the sentence itself suggests and real world knowledge, iv) find ways to represent them through BNs, and v) make inferences.

Let us go back to the research questions posed in Chapter 1, and answer them in light of what we have laid out with this approach:

1. Does the antecedent need to be false or unrealized for one to be able to use a counterfactual construction?

No. We have seen in Finding the Medicine example that one's lack of knowledge or uncertainty can lead one to use a counterfactual construction. There exist additional instances of non-counterfactual statements that possess a counterfactual character. A similar example from English would be (85), adapted from Anderson (1951).

(85) [If Ali had ingested cyanide], he would show the precise symptoms he is in fact showing.
[Ali siyanür almış/içmiş/tüketmiş olsaydı] tam da şu anda gösterdiği semptomları gösterirdi.

It is clear that an individual who espouses the belief that Ali has indeed ingested cyanide may appropriately utter (85) without designating the antecedent as counterfactual. This underscores the complexity and variability of counterfactual reasoning.

2. Does the use of *-sAydI*, a complex suffix that is a combination of the conditional marking verbal suffix *-sA* and past tense marking copula *-(y)dI*, in a Turkish sentence always lead to a counterfactual interpretation?

No; again, we have seen in Finding the Medicine example that one can use *-sAydI* without having a definite counterfactuality in mind.

The morphological marking that differentiates conditionals referred to as counterfactual from those that are not, can also be observed in other modal constructions such as the expression of wishes and obligations. Iatridou (2000) calls instances of past morphology that do not convey a temporal past interpretation as *fake tense*. von Fintel & Iatridou (2023) call this *X-marking*; they present the requirements for a comprehensive theory of X-marking and provide some preliminary informal observations, and according to them, Turkish is a *fake past language*.

(42)[19] is repeated below, in which verbs of both the antecedent and the consequent are marked with a *fake* past.

(42) [Yapacak çok işim olma**saydı**] ben de yarınki partiye gelir**di**m.
I would have come to the party tomorrow, too, [if I had not had a lot to do].

Note that this use of -*sAydI* with a certain future reference serves to highlight one of the shortcomings of the approach we proposed in that there is not a well-defined way to represent such a scenario.

3. What happens when -*DI* precedes -*(y)sA* and what happens when it is vice versa? In other words, what exactly is the difference between -*DIysA* and -*sAydI*?

We mentioned in Chapter 4 that the classification of conditional statements as counterfactual or hypothetical is an ongoing topic of debate among scholars. According to Ruhi et al. (2000), the version with the conditional marker at the beginning, -sAydI, is indicative of a counterfactual statement, while its positioning at the end, -DIysA, is indicative of a hypothetical statement. Can Bakırlı (2010) and Oktar & Can (2016) share this same analysis. In contrast, Yarbay Duman et al. (2016) designate the former as counterfactual and the latter as factual in their study, although their analysis is not entirely dissimilar from the aforementioned scholars. All in all, we argue for the inclusion of and the need to focus on *context* in these discussions.

Additionally, in 4.2.1, we posited that the differentiation between -*DIysA* and -*sAydI* is not restricted solely to syntax and semantics. Instead, this differentiation has event-referential implications, whereby they align with different levels in Pearl's framework. Moreover, this differentiation has representational consequences, as the event-referential difference necessitates these constructions to be depicted through two distinct BNs.

4. On which rung of Pearl's Ladder of Causation do the sentences which involve the use -*DIysA* and -*sAydI* fall?

In 4.2.1, we put forward the idea that -*DIysA* statements are related to observations, and they fall on rung 1; aorist + -*(y)sA* statements are related to interventions, and they correspond to rung 2; and -*sAydI* statements are related to counterfactuals which means they fall on rung 3.

5. Does pragmatics, or to be more precise, context play a role in the interpretation of a counterfactual statement?

[19] This type of sentences is called Future Less Vivid, see Iatridou (2000) for details.

We have shown at least two ways how context can affect the interpretation of a counterfactual statement in Finding the Medicine example. Let us expand upon that scenario with a small addition that is context-dependent.

Consider (86), the first person alternatives of the medicine sentence.

(86) [İlacı bulmasaydım] hastalığım ilerleyecekti.

[If I had not found the medicine], my disease would have worsened.

Since it is a first person account of what would have happened, it does not seem like adding *bildiğim kadarıyla* (*as far as I know/to the best of my knowledge*) sentence-initially as in or *ne yaptı acaba* (*I wonder what they did*) sentence-finally as we did earlier for the third person account is reasonable; however, when you consider a context in which this person was in a coma because of the said disease, a sentence-final addition such as acaba *ne oldu da iyileştim* (*I wonder what it was that happened and allowed me to get better*) becomes much more plausible; so, context is essential in analyzing counterfactuals as even one's own account of oneself can incorporate uncertainty.

CHAPTER 5

5. CONCLUSION & FUTURE WORK

If we didn’t have birthdays, you wouldn’t be you.
If you’d never been born, well then what would you do?
If you’d never been born, well then what would you be?
You might be a fish! Or a toad in a tree!
You might be a doorknob! Or three baked potatoes!
You might be a bag full of hard green tomatoes.
Or worse than all that . . . You might be a WASN’T!
A Wasn’t has no fun at all. No, he doesn’t.
A Wasn’t just isn’t. He just isn’t present.
But you . . . You are YOU! And, now isn’t that pleasant.

Dr. Seuss - Happy Birthday to You! (1959)

5.1. Conclusion

This thesis aimed to analyze counterfactual statements in Turkish by combining linguistic and causal modeling perspectives, which was a novel approach that provided new insights into their nature. Counterfactuals are generally defined as conditional statements with false or unrealized antecedents, and this thesis challenged this definition.

In Chapter 1, the concept of counterfactuals was introduced, and their use in different academic fields was discussed. Chapter 2 presented a literature review on conditionals and counterfactuals, with a particular focus on Turkish conditionals and counterfactuals. Chapter 3 discussed causality and causal inference, including an introduction to the Ladder of Causation and an in-depth examination of its three rungs. It also introduced the graphical notation and terminology used for causal inferences and demonstrated their application in representing different kinds of inference contexts. Finally, Chapter 4 presented arguments against the assertion that counterfactuals were conditionals with false or unrealized antecedents, highlighted the significance of context in interpreting counterfactuals, and introduced an approach that enabled various scenarios to be represented.

The approach we have developed and delineated provided a relatively robust methodology for representing a wide range of observation, intervention, and counterfactual scenarios, as well as scenarios with uncertain evidence and/or unexpected, exogenous factors. In our analysis, we not only consider the state of affairs described in the antecedent-consequent duo, but also enrich its context, define the causal relationships between the products of the said enrichment, assign arbitrary probabilities to them based on what the sentence itself suggests and real-world knowledge, find ways to represent them through BNs, and make inferences.

With this approach in mind, we revisited the research questions posed in Chapter 1 and provided some answers. First, we have seen that the antecedent need not be false or unrealized for one to use a counterfactual construction. As evidenced by the Finding the Medicine example and the example we adapted from Anderson's (1951), individuals may use counterfactual constructions due to a lack of knowledge or uncertainty. Additionally, there exist non-counterfactual statements that possess a counterfactual character, highlighting the complexity and variability of counterfactual reasoning.

Second, the use of the complex suffix *-sAydI* in a Turkish sentence does not always lead to a counterfactual interpretation. Again, the Finding the Medicine example demonstrates that individuals may use *-sAydI* without having a definite counterfactual in mind. Furthermore, the morphological marking that differentiates counterfactual from non-counterfactual conditionals can also be observed in other modal constructions such as the expression of wishes and obligations.

Third, we have explored the differences between the two Turkish constructions *-DIysA* and *-sAydI*. *-sAydI* is indicative of a counterfactual statement while *-DIysA* is indicative of a hypothetical statement. Our analysis suggests that context plays a significant role in the differentiation between these constructions and their event-referential implications. Specifically, *-DIysA* statements are related to observations and fall on rung 1 of Pearl's Ladder of Causation, aorist + *-(y)sA* statements are related to interventions and correspond to rung 2, and *-sAydI* statements are related to counterfactuals and fall on rung 3.

Finally, we have shown that pragmatics and context play a crucial role in the interpretation of a counterfactual statement. Our analysis of the Finding the Medicine example demonstrated that the same sentence can be interpreted differently based on the interlocutors' knowledge and assumptions. This finding highlights the importance of considering the broader context when analyzing counterfactual statements.

In conclusion, we have shown through bringing together two different approaches to counterfactuals that the context is simply essential to any analysis. Our interdisciplinary approach to counterfactuals underscores the value of integrating diverse perspectives to deepen our understanding of complex cognitive phenomena, and we hope that our findings will inspire further research in this area.

5.2. Future Work

This subchapter lists what can be done further regarding this topic.

5.2.1. Introspective Judgments

The validity of statements that are derived through introspective judgments (as is the case in most of the linguistics studies cited in this work) remains a topic of debate. To reinforce the theoretical claims, an experimental approach could be employed. Such an experiment may incorporate a Cloze task in which participants are required to provide consequents for various *-sAydI* and *-DIysA* antecedents in different contexts. Additionally, an acceptability judgment task could also be implemented (see Gibson & Fedorenko, 2010, 2013 on the need for quantitative methods in linguistics research).

5.2.2. Fake Tense & X-Marking

We were only able to barely touch upon the now to-be-published von Fintel & Iatridou (2023) piece on what they call X-marking, exploring it further and reviewing other works on the same topic, such as Mackay (2015), and Schulz (2014) are endeavors that are sure to prove fruitful.

5.2.3. Negation

We have seen in 4.3 that how the counterfactuals interact with negation is an interesting topic and made an early prediction that it will have something to do with conditional perfection (strengthening of a conditional to biconditional). The works that can be investigated regarding this include but are not limited to: Espino et al. (2018, 2022), Ferguson et al. (2008), Gómez-Sánchez et al. (2021), Horn (2000), Ippolito & Su (2014), Orenes et al. (2022), Pfeifer & Yama (2017), and Schulz (2018).

5.2.4. Lexical Aspect

The concept of lexical aspect is a fundamental aspect of natural language semantics and its relevance for explaining a wide array of linguistic phenomena has been well-established. It intersects with various linguistic concepts such as grammatical aspect, tense, adverbial modification, the syntax and semantics of quantification and expressions of quantity, argument structure, linking at the lexical semantics-syntax interface, and plays a role in the temporal sequencing of discourse. Lexical aspect is a semantic category that pertains to the properties of eventualities, as defined by Bach (1981), as expressed by verbs. In general terms, the properties in question pertain to the presence or absence of an endpoint, limit, or boundary within the lexical structure of certain classes of verbs (Filip, 2012). We believe there is an interaction between a verb's lexical aspect and behavior in a counterfactual context. The works that can be investigated regarding this include but are not limited to: Comrie (1976), Güven (2012), Han (1996), Moens & Steedman (1988), Smith (1997), and Vendler (1957).

5.2.5. Other Related Works

There have been other works that brought together linguistics, causality, probabilistic reasoning, and counterfactuals to some extent. Some of the recent ones that can potentially add a great deal to our discussion of counterfactuals include but are not limited to:

Lassiter (2017a) addresses an issue in theories of counterfactuals based on maximal similarity or minimal revision involving negated conjunctions in the antecedent. The article proposes a theory that implements general probabilistic causal knowledge when there are multiple ways of instantiating a counterfactual antecedent.

Lassiter (2017b, 2018[20]) provides a unified compositional semantics for indicative and counterfactual conditionals with probability operators. The article combines a Kratzerian syntax for conditionals with semantics based on BNs and develops an algorithm for interpreting probabilistic indicatives and counterfactuals.

Lassiter (2020) presents an approach that uses a simple probabilistic model to handle problems of uncertainty and ignorance, taking into account the hierarchical structure of information states and using BNs. The article also addresses the interaction between probability operators and counterfactuals and proposes two extensions of the semantics that resolve the resulting puzzle.

Over (2017) explores the relationship between natural language conditionals, probability judgments, and causation. The article discusses the role of causal conditionals in explaining causation, and the connection between Bayesian inferences, conditional probability, and imaging.

Popa (2022) examines the relationship between philosophical and psychological research on causation and counterfactual thought, with a focus on the problem of backtracking. While some models of counterfactual causation rule out backtracking, various formal models and psychological evidence suggest that people sometimes use backtracking counterfactuals. The article argues that non-backtracking counterfactuals are more commonly used in causal inference because they are consistent with temporal order information and match reasoners' experience of causation.

---

[20] Lassiter (2018) is an updated version of Lassiter (2017b) but it is worth examining both of them.

REFERENCES

Adams, E. W. (1970). Subjunctive and Indicative Conditionals. *Foundations of Language*, *6*(1), 89–94.

Anderson, A. R. (1951). A Note on Subjunctive and Counterfactual Conditionals. *Analysis*, *12*(2), 35–38.

Athanasiadou, A. (1997). Introduction. A. Athanasiadou, & R. Dirven (Editors), *On Conditionals Again*, 1–12.

Austin, J. L. (1956). Ifs and Cans. *Proceedings of the British Academy*, *42*, 109–132.

Austin, J. L. (1975). *How to Do Things with Words: The William James Lectures Delivered at Harvard University in 1955.*

Bach, E. 1981. On Time, Tense and Aspect: An Essay in English Metaphysics. P. Cole (editor), *Radical Pragmatics*, 63–81.

Balke, A., & Pearl, J. (1994). Probabilistic Evaluation of Counterfactual Queries. *Proceedings of the Twelfth National Conference on Artificial Intelligence*, *1*, 230–237.

Barwise, J. (1986). Conditionals and Conditional Information. E. Traugott, A. Meulen, J. Reilly, & C. Ferguson (Authors), *On Conditionals*, 21–54.

BayesFusion. (2022). GeNIe Academic Version 4.0.2423. https://www.bayesfusion.com/genie

Bernanke, B. S., Gertler, M., & Gilchrist, S. (1999). The Financial Accelerator in a Quantitative Business Cycle Framework. *Handbook of Macroeconomics*, *1*, 1341–1393.

Bhatt, R., & Pancheva, R. (2017). Conditionals. *The Wiley Blackwell Companion to Syntax, Second Edition*, 1–48.

Byrne, R. M. (2016). Counterfactual Thought. *Annual Review of Psychology*, *67*(1), 135–157.

Can Bakırlı, Ö. (2010). *Türkçede Koşullu Yapılar*. Doctoral Dissertation, DEÜ Sosyal Bilimler Enstitüsü.

Card, D., & Krueger, A. B. (1995). *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton University Press.

Comrie, B. (1976). *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge University Press.

Comrie, B. (1986). Conditionals: A Typology. E. Traugott, A. Meulen, J. Reilly, & C. Ferguson (Authors), *On Conditionals*, 77–100.

Culicover, P. W., & Jackendoff, R. (1999). The View from the Periphery: The English Comparative Correlative. *Linguistic Inquiry*, *30*(4), 543–571.

Declerck, R., & Reed, S. (2012). *Conditionals: A Comprehensive Empirical Analysis*. Walter de Gruyter.

Edgington, D. (2020). Indicative Conditionals. E. N. Zalta (editor), *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition).

Egré, P., & Rott, H. (2021) The Logic of Conditionals. E. N. Zalta (editor), *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition).

Ehrlich, I. (1996). Crime, Punishment, and the Market for Offenses. *The Journal of Economic Perspectives*, *10*(1), 43–67.

Epstude, K., & Roese, N. J. (2008). The Functional Theory of Counterfactual Thinking. *Personality and Social Psychology Review*, *12*(2), 168–192.

Espino, O., & Byrne, R. M. (2018). Thinking About the Opposite of What Is Said: Counterfactual Conditionals and Symbolic or Alternate Simulations of Negation. *Cognitive Science*, *42*(8), 2459–2501.

Espino, O., Orenes, I., & Moreno-Ríos, S. (2022). Inferences from the Negation of Counterfactual and Semifactual Conditionals. *Memory & Cognition*, *50*(5), 1090–1102.

Feinberg, J. (1970). *Doing & Deserving; Essays in the Theory of Responsibility*. Princeton University Press.

Ferguson, H. J., Sanford, A. J., & Leuthold, H. (2008). Eye-Movements and Erps Reveal the Time Course of Processing Negation and Remitting Counterfactual Worlds. *Brain Research, 1236*, 113–125.

Ferguson, N. (1997). *Virtual History: Alternatives and Counterfactuals*. Picador.

Filip, H. (2012). Lexical Aspect. Robert I. Binnick (editor), *The Oxford Handbook of Tense and Aspect*, 721–751.

von Fintel, K., & Iatridou, S. (2023). Prolegomena to a theory of X-marking. Final version, to appear in *Linguistics and Philosophy*.

von Fintel, K. (1994). *Restrictions on Quantifier Domains*. Doctoral Dissertation, University of Massachusetts at Amherst.

von Fintel, K. (2011). Conditionals. K. von Heusinger, C. Maienborn, & P. Portner (Editors), *Semantics: An International Handbook of Meaning*, *2*, 1515–1538.

von Fintel, Kai. (2012). Subjunctive conditionals. G. Russell, & D. R. Graff Fara (Editors), *The Routledge Companion to Philosophy of Language*, 466–477.

Gibson, E., & Fedorenko, E. (2010). Weak Quantitative Standards in Linguistics Research. *Trends in Cognitive Sciences*, *14*(6), 233–234.

Gibson, E., & Fedorenko, E. (2013). The Need For Quantitative Methods In Syntax And Semantics Research. *Language and Cognitive Processes*, *28*(1–2), 88–124.

Göksel, A., & Kerslake, C. (2005). *Turkish: A Comprehensive Grammar*. Routledge.

Gómez-Sánchez, J., Moreno-Ríos, S., & Frosch, C. (2021). Alternatives or Syntactic Negation? Adults' and Children's Preferences for Constructing Counterfactual Possibilities. *Current Psychology*, 1–13.

Greenberg, J. H. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. *Universals of Language*, *2*, 73–113.

Güven, M. (2012). On the Linguistic Realization of Situation Types in Turkish. *Trakya Üniversitesi Sosyal Bilimler Dergisi*, *14*(1), 177–194.

Han, C. H. (1996). Comparing English and Korean Counterfactuals: The Role of Verbal Morphology and Lexical Aspect in Counterfactual Interpretation. *Proceedings of Eastern States Conference on Linguistics*, *96*, 124–137.

Hanushek, E. A. (2011). The Economic Value of Higher Teacher Quality. *Economics Of Education Review*, *30*(3), 466–479.

Horn, L. R. (2000). From *if* to *iff*: Conditional Perfection as Pragmatic Strengthening. *Journal of Pragmatics*, *32*(3), 289–326.

Iatridou, S. (2000). The Grammatical Ingredients of Counterfactuality. *Linguistic Inquiry*, *31*(2), 231–270.

Ippolito, M., & Su, J. (2014). Counterfactuals, Negation, and Polarity. *The Art and Craft of Semantics: A Festschrift for Irene Heim*, *1*, 225–43.

JGraph. (2023). draw.io / diagrams.net Flowchart Maker and Online Diagram Software 20.8.16. https://www.drawio.com

Kahneman, D., & Tversky, A. (1982). The Simulation Heuristic. D. Kahneman, P. Slovic, & A. Tversky (Editors), *Judgment under Uncertainty: Heuristics and Biases*, 201–208.

Kratzer, A. (1981). The Notional Category of Modality. H. Eikmeyer & H. Rieser (Editors), *Words, Worlds, and Contexts: New Approaches in Word Semantics*, 38–74.

Kusner, M. J., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual Fairness. I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Editors), *Advances in Neural Information Processing Systems*, *30*.

Lassiter, D. (2017a). Complex Antecedents and Probabilities in Causal Counterfactuals. *21st Amsterdam Colloquium*, 45–54.

Lassiter, D. (2017b). Probabilistic Language in Indicative and Counterfactual Conditionals. *Semantics and Linguistic Theory*, *27*, 525–546.

Lassiter, D. (2018). Bayes Nets and the Dynamics of Probabilistic Language. *Proceedings of Sinn und Bedeutung*, *21*(2), 747–766.

Lassiter, D. (2020). *Causation and Probability in Indicative and Counterfactual Conditionals*.

Lauritzen, S. L., & Spiegelhalter, D. J. (1988). Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, *50*(2), 157–224.

Lewis, D. (1973a). *Counterfactuals*. Blackwell.

Lewis, D. (1973b). Counterfactuals and Comparative Possibility. *Journal of Philosophical Logic*, *2*, 418–446.

Mackay, J. (2015). Actuality and Fake Tense in Conditionals. *Semantics and Pragmatics*, *8*, 1–12.

Mackie, J. L. (1965). Causes and Conditions. *American Philosophical Quarterly*, *2*(4), 245–264.

Manski, C. F. (2008). *Identification for Prediction and Decision*. Harvard University Press.

Marcus, R. B. (1961). Modalities and Intensional Languages. *Synthese*, *13*(4), 303–322.

Moens, M., & Steedman, M. (1988). Temporal Ontology and Temporal Reference. *Computational Linguistics*, *14*(2), 15–28.

Mrad, A. B., Delcroix, V., Piechowiak, S., Leicester, P., & Abid, M. (2015). An Explication of Uncertain Evidence in Bayesian Networks: Likelihood Evidence and Probabilistic Evidence. *Applied Intelligence*, *43*, 802–824.

Orenes, I., Espino, O., & Byrne, R. M. (2022). Similarities and Differences in Understanding Negative and Affirmative Counterfactuals and Causal Assertions: Evidence from Eye-Tracking. *Quarterly Journal of Experimental Psychology*, *75*(4), 633–651.

Over, D. E. (2017). Causation and the Probability of Causal Conditionals. M. R. Waldmann (editor), *The Oxford Handbook of Causal Reasoning*, 307–325.

Pearl, J. (1993). Aspects of Graphical Models Connected with Causality. *Proceedings of the 49th Session of the International Statistical Institute*, 399–401.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books.

Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal Inference in Statistics: A Primer*. Wiley.

Pfeifer, N., & Yama, H. (2017). Counterfactuals, Indicative Conditionals, and Negation under Uncertainty: Are There Cross-Cultural Differences? *arXiv:1703.03255*

Plantinga, A. (1974). *The Nature of Necessity*. Oxford University Press.

Popa, E. (2022). Getting Counterfactuals Right: The Perspective of the Causal Reasoner. *Synthese*, *200*(1), 17.

Ramsey, F. P. (1931). General Propositions and Causality. R. B. Braithwaite (editor), *The Foundations of Mathematics and Other Logical Essays*, 237–255.

Roese, N. J., & Olson, J. M. (1995). Counterfactual Thinking: A Critical Overview. N. J. Roese, & J. M. Olson (Editors), *What Might Have Been: The Social Psychology of Counterfactual Thinking*, 1–55.

Ruhi, Ş., Zeyrek, D., & Turan, Ü. D. (2000). Koşul Tümcelerinde Varsayımsallık ve Gerçek Karşıtlığı. A. S. Özsoy & E. E. Taylan (Editors), *XIII. Dilbilim Kurultayı Bildirileri*, 19–29.

Schreiber, J. (2018). pomegranate: Fast and Flexible Probabilistic Modeling in Python. *Journal of Machine Learning Research*, *18*, 1–6.

Schulz, K. (2014). Fake Tense in Conditional Sentences: A Modal Approach. *Natural Language Semantics*, *22*(2), 117–144.

Schulz, K. (2018). The Similarity Approach Strikes Back: Negation in Counterfactuals. *Proceedings of Sinn und Bedeutung*, *22*(2), 343–360.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.

Shpitser, I., & Pearl, J. (2008). Complete Identification Methods for the Causal Hierarchy. *Journal of Machine Learning Research*, *9*, 1941–1979.

Smith, C. S. (1997). *The Parameter of Aspect*. Kluwer Academic Publishers.

Starr, W. (2022). Counterfactuals. E. N. Zalta, & U. Nodelman (Editors), *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition).

Russell, S.J., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. Pearson Education.

Stump, G. T. (1985). *The Semantic Variability of Absolute Constructions*. D. Reidel Publishing Company.

Swaminathan, A., & Joachims, T. (2015). The Self-Normalized Estimator for Counterfactual Learning. C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Editors), *Advances in Neural Information Processing Systems*, *28*.

Thompson, S., Longacre, R., & Hwang, S. (2007). Adverbial Clauses. T. Shopen (Editor), *Language Typology and Syntactic Description*, 237–300.

Vendler, Z. (1957). Verbs and Times. *The Philosophical Review*, *66*(2), 143–160.

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Yarbay Duman, T., Altınok, N., & Maviş, İ. (2016). Grammar and Cognition: Deficits Comprehending Counterfactuals in Turkish Individuals with Broca's Aphasia. *Aphasiology*, *30*(7), 841–861.

Yarbay Duman, T., Blom, E., & Topbaş, S. (2015). At the Intersection of Cognition and Grammar: Deficits Comprehending Counterfactuals in Turkish Children with Specific Language Impairment. *Journal of Speech, Language, and Hearing Research*, *58*(2), 410–421.

Yu, E. (2021). Bayesian Neural Networks with Soft Evidence. *arXiv:2010.09570*