Enabling Social Situation Awareness in Support Agents

Kola, I.

DOI
10.4233/uuid:46bc9dd8-c8a9-4de4-ab18-9b53b700e4bd

Publication date
2022

Document Version
Final published version

Citation (APA)
Kola, I. (2022). *Enabling Social Situation Awareness in Support Agents*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:46bc9dd8-c8a9-4de4-ab18-9b53b700e4bd

Enabling Social Situation Awareness in Support Agents

Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology,
by the authority of the Rector Magnificus Prof. dr.ir. T.H.J.J. van der Hagen,
Chair of the Board for Doctorates,
to be defended publicly on
Monday 21 November 2022 at 15:00 o'clock

by

Ilir KOLA

Master of Science in Computer Science,
University of Freiburg, Germany,
born in Peshkopi, Albania.

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus,	Chairperson
Prof. dr. C.M. Jonker,	Delft University of Technology, promotor
Dr. M.B. Van Riemsdijk	University of Twente, copromotor
Independent Members:	
Prof. dr. M. A. Neerincx	Delft University of Technology
Prof. dr. D. K. J. Heylen	University of Twente
Prof. dr. P. Yolum	University of Utrecht
Dr. J. Cassens	University of Hildesheim, Germany
Prof. dr. M. E. Warnier	Delft University of Technology, reserve member

SIKS Dissertation Series No. 2022-23
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

This thesis is part of the research programme CoreSAEP, with project number 639.022.416, which is financed by the Netherlands Organisation for Scientific Research (NWO).

Printed by: ProefschriftMaken (www.proefschriftmaken.nl)

Front & Back: Simone Golob (www.sgiv.nl)

ISBN 978-94-6469-099-6
An electronic version of this dissertation is available at
http://repository.tudelft.nl/.

Dedicated to my dad,
who taught me to love life, knowledge, and nature.

Contents

SUMMARY

The use of support agents that help people in their daily lives is steadily growing. While there have been continuous developments in integrating and modelling internal aspects of the user in these support agents, research shows that people's behavior is also shaped by their environment. Currently there have been attempts at integrating elements of the physical environment such as location, however, support agents generally lack the ability to take into account the effect of the user's social situation on their behavior. This is important since the majority of our daily life situations have a social nature.

This thesis proposes a *social situation awareness* framework for allowing support agents to take into account the user's social situation in order to offer more comprehensive support. This is based on our definition of social situation awareness in support agents: 'a support agent's ability to perceive the social elements of a situation, to comprehend their meaning, and to infer their effect on the behavior of the user'. The framework is inspired by existing work on situation awareness from research in human factors and computer science, instantiated with concepts from social sciences.

First, we propose a conceptual architecture for social situation awareness in support agents. The architecture is based on a proposed set of requirements that need to be addressed in order for support agents to exhibit social situation awareness. The core component of the architecture is the social situation awareness module consisting of three levels: 1) *social situation perception*, in which the social background features and social situation cues are perceived and modelled; 2) *social situation comprehension*, in which meaning is ascribed to the situation in the form of a social situation profile; 3) *social situation projection*, in which the expected user behavior and the affected personal values are assessed. Furthermore, the architecture contains several interactive modules responsible for eliciting information from users, support, explanations and feedback.

The rest of the thesis focuses on implementing the three levels of situation awareness and allowing support agents to transition between them. To do so, we combine concepts from social science related to social relationships, psychological characteristics of situations and personal values with machine learning techniques. Throughout the thesis we illustrate our proposed approach through a hypothetical case study evolving around a socially aware agenda management agent that helps users deal with overlapping meetings. This case study is suitable as a proof of concept for our approach since meetings have an inherently social nature, allowing us to explore how we can incorporate the influence of the social aspects of situations in support agents.

To realize social situation perception, we propose a two-level ontology for modelling social situations. The ontology focuses on modelling the social relationships between the user and other people, as well as other features of the social situation. The modelled concepts are based on existing work on social relationships and situations. We propose a set of social situation features that can be used to model meetings between the user and another person, and the features are validated through a user study. To further evaluate the usefulness of these features, we conduct a larger crowd-sourcing user study where participants provided information about people from their social circle using our proposed features, and then answer questions about the priority level that they would assign to hypothetical situations involving those people. The data is used to build machine learning models that take as input the social situation features and predict the priority level of social situations. Results show that such a model achieves a significantly better prediction accuracy than baseline predictors.

Next, we focus on how to achieve social situation comprehension in support agents. We propose that psychological characteristics of situations, a concept researched in social psychology, can serve as a basis for ascribing meaning to situations in support agents. Particularly, we focus on the existing DIAMONDS taxonomy, which consists of the psychological characteristics Duty, Intellect, Adversity, Mating, Positivity, Negativity, Deception and Sociality. Using a data set containing information about a large number of meeting situations annotated in terms of their social situation features, psychological characteristics, and priority level, we explore how social situation comprehension can be implemented in support agents. First of all, we demonstrate that psychological characteristics of situations are a better predictor of the priority of social situations than social situation features. Then, we explore how we can use machine learning models to predict the psychological characteristics of situations using as input social situation features. These predictive models show that it is possible to transition between the levels of the architecture: Level 2 information can be predicted from Level 1 information, and Level 3 information can be predicted from Level 2 information.

Personal values are considered key drivers of human behavior, therefore they are crucial for support agents that need to assess the user's behavior. Values are situational, therefore apart from information regarding which values are important to the user, the agent also needs to know how social situations affect values. Specifically, we explore whether a social situation promotes, demotes, or does not affect the ability of the user to fulfill a personal value. This information would allow support agents to provide value-aligned support. Assessing how situations affect personal values is part of Level 3 - social situation projection. To represent personal values, we use a validated list from existing research in social sciences. We conduct a user study in which participants describe social situations from their daily life, and answer questions about the psychological characteristics of these situations as well as about whether these situations promote, demote or do not affect a set of personal values. We group similar situations based on their psychological characteristics, and show that specific situation groups significantly promote or demote certain personal values. For instance, situations with a high level of duty and intellect promote values such as helpfulness and capability, while situations with a high level of adversity demote the value safety.

This thesis demonstrates how to integrate social situation awareness components in support agents. The studies presented in the thesis provide insight into the concepts and techniques that are needed for social situation awareness, and how they can be used in practice through a hypothetical case study involving a socially aware agenda management agent. This contribution serves as a blueprint for designers of support agents, and provides a basis towards more comprehensive support for users.

Samenvatting

Het gebruik van ondersteuningsagenten die mensen helpen in hun dagelijks leven neemt gestaag toe. Hoewel er voortdurend ontwikkelingen zijn geweest in het integreren en modelleren van interne aspecten van de gebruiker in deze ondersteuningsagenten, blijkt uit onderzoek dat het gedrag van mensen ook wordt gevormd door hun omgeving. Hoewel er pogingen zijn gedaan om elementen van de fysieke omgeving, zoals de locatie, te integreren, zijn ondersteuningsagenten over het algemeen niet in staat om rekening te houden met het effect van de sociale situatie van de gebruiker op zijn gedrag. Dit is belangrijk omdat de meeste situaties in ons dagelijks leven een sociaal karakter hebben.

In dit proefschrift wordt een raamwerk voor *sociale situatiebewustzijn* voorgesteld om ondersteuningsagenten in staat te stellen rekening te houden met de sociale situatie van de gebruiker om meer uitgebreide ondersteuning te bieden. Dit is gebaseerd op onze definitie van sociale situatiebewustzijn in ondersteuningsagenten: 'het vermogen van een ondersteuningsagent om de sociale elementen van een situatie waar te nemen, om hun betekenis te begrijpen en om hun effect op het gedrag van de gebruiker af te leiden'. Het raamwerk is geïnspireerd door bestaand werk over situatiebewustzijn uit onderzoek naar menselijke factoren en uit de informatica, aangevuld met concepten uit de sociale wetenschappen.

Ten eerste introduceren we een conceptuele architectuur voor sociale situatiebewustzijn in ondersteuningsagenten. De architectuur is gebaseerd op een voorgestelde verzameling van vereisten waaraan moet worden voldaan zodat ondersteuningsagenten sociale situatiebewustzijn vertonen. De kerncomponent van de architectuur is de sociale situatiebewustzijnsmodule die bestaat uit drie niveaus: 1) *sociale situatieperceptie*, waarin de sociale achtergrondkenmerken en sociale situatiecues worden waargenomen en gemodelleerd; 2) *sociale situatiebegrip*, waarin betekenis wordt toegekend aan de situatie in de vorm van een sociale situatieprofiel; 3) *sociale situatieprojectie*, waarin het verwachte gebruikersgedrag en de beïnvloede persoonlijke waarden worden beoordeeld. Verder bevat de architectuur verschillende interactieve modules die verantwoordelijk zijn voor het uitvragen van informatie bij gebruikers, voor ondersteuning, en voor uitleg en feedback.

De rest van het proefschrift richt zich op het implementeren van de drie niveaus van situatiebewustzijn en het mogelijk maken voor ondersteuneningsagenten om tussen deze niveaus te schakelen. Daartoe combineren wij concepten uit de sociale wetenschappen met betrekking tot sociale relaties, psychologische kenmerken van situaties en persoonlijke waarden met technieken voor machinaal leren. In dit proefschrift illustreren we onze voorgestelde aanpak aan de hand van een hypothetische casus rond een sociaal bewuste agendabeheeragent die gebruikers helpt om te gaan met overlappende vergaderingen. Deze casus is geschikt als proof of concept voor onze aanpak, aangezien vergaderingen een inherent sociaal karakter hebben, waardoor we kunnen onderzoeken

hoe we de invloed van de sociale aspecten van situaties kunnen opnemen in ondersteuneningsagenten.

Om sociale situatieperceptie te realiseren, stellen wij een ontologie op twee niveaus voor om sociale situaties te modelleren. De ontologie richt zich op het modelleren van de sociale relatie tussen de gebruiker en andere mensen, alsmede op andere kenmerken van de sociale situatie. De gemodelleerde concepten zijn gebaseerd op bestaand werk op het gebied van sociale relaties en situaties. Wij stellen een verzameling van sociale situatiekenmerken voor die kunnen worden gebruikt om ontmoetingen tussen de gebruiker en een andere persoon te modelleren, en de kenmerken zijn gevalideerd aan de hand van een gebruikersstudie. Om de bruikbaarheid van deze kenmerken verder te evalueren, voeren we een groter crowdsourcing gebruikersonderzoek uit waarbij deelnemers eerst informatie verstrekken over mensen uit hun sociale kring aan de hand van de door ons voorgestelde kenmerken, en vervolgens beantwoorden ze vragen over het prioriteitsniveau dat zij zouden toekennen aan hypothetische situaties waarbij die mensen betrokken zijn. De gegevens worden gebruikt om machinale leermodellen te bouwen die de sociale situatiekenmerken als input nemen en het prioriteitsniveau van sociale situaties voorspellen. Uit de resultaten blijkt dat een dergelijk model een aanzienlijk betere voorspellingsnauwkeurigheid bereikt dan modellen die geen gebruik maken van deze gegevens.

Vervolgens richten we ons op de vraag hoe sociale situatiebegrip in ondersteuningsagenten kan worden bereikt. Wij stellen voor dat psychologische situatiekenmerken, een concept dat in de sociale psychologie is onderzocht, kunnen dienen als basis voor het toekennen van betekenis aan situaties in ondersteuningsagenten. In het bijzonder richten wij ons op de bestaande DIAMONDS taxonomie, die bestaat uit de psychologische kenmerken Plicht, Intellect, Tegenspoed, Paring, Positiviteit, Negativiteit, Bedrog en Socialiteit. Aan de hand van een dataset met informatie over een groot aantal ontmoetingssituaties, geannoteerd in termen van hun sociale situatiekenmerken, psychologische kenmerken en prioriteitsniveau, onderzoeken wij hoe het begrip van sociale situaties kan worden geïmplementeerd in ondersteuningsagenten. Allereerst tonen we aan dat psychologische situatiekenmerken een betere voorspeller zijn van de prioriteit van sociale situaties dan sociale situatiekenmerken. Vervolgens onderzoeken we hoe we machinale leermodellen kunnen gebruiken om de psychologische situatiekenmerken te voorspellen met als input sociale situatiekenmerken. Deze voorspellende modellen laten zien dat het mogelijk is om tussen de niveaus van de architectuur te schakelen: Niveau 2 informatie kan worden voorspeld uit Niveau 1 informatie, en Niveau 3 informatie kan worden voorspeld uit Niveau 2 informatie.

Persoonlijke waarden worden beschouwd als belangrijke drijfveren van menselijk gedrag, en zijn daarom cruciaal voor ondersteuningsagenten die het gedrag van de gebruiker moeten inschatten. Waarden zijn situationeel, dus behalve informatie over welke waarden belangrijk zijn voor de gebruiker, moet de agent ook weten hoe sociale situaties waarden beïnvloeden. Specifiek onderzoeken we of een sociale situatie het vermogen van de gebruiker om een persoonlijke waarde te vervullen bevordert, belemmert of niet beïnvloedt. Met deze informatie zouden ondersteuningsagenten op waarden afgestemde ondersteuning kunnen bieden. Inschatten hoe situaties persoonlijke waarden beïnvloeden maakt deel uit van Niveau 3 - projectie van sociale situaties. Om persoonlijke waarden te representeren, gebruiken we een gevalideerde lijst uit bestaand onderzoek in de sociale wetenschappen. Wij voeren een gebruikersonderzoek uit waarin deelnemers sociale situaties uit hun dagelijks leven beschrijven, en vragen beantwoorden over de psychologische kenmerken van deze situaties en over de vraag of deze situaties een set van persoonlijke waarden bevorderen, belemmeren of niet beïnvloeden. Wij groeperen vergelijkbare situaties op basis van hun psychologische kenmerken, en tonen aan dat specifieke situatiegroepen bepaalde persoonlijke waarden significant bevorderen of belemmeren. Zo bevorderen situaties met een hoge mate van plichtsbesef en intellect waarden als hulpvaardigheid en bekwaamheid, terwijl situaties met een hoge mate van tegenspoed de waarde veiligheid belemmeren.

Dit proefschrift laat zien hoe sociale situatiebewustzijn kan worden geïntegreerd in ondersteuningsagenten. De in dit proefschrift gepresenteerde studies geven inzicht in de concepten en technieken die nodig zijn voor sociale situatiebewustzijn, en hoe deze in de praktijk kunnen worden gebruikt aan de hand van een hypothetische casestudie met een sociaal bewuste agendabeheeragent. Deze bijdrage dient als blauwdruk voor ontwerpers van ondersteuningsagenten, en biedt een basis voor meer uitgebreide ondersteuning voor gebruikers.

PËRMBLEDHJE

Përdorimi i agjentëve të suportit[1] që ndihmojnë njerëzit në përditshmërinë e tyre po rritet në mënyrë të vazhdueshme. Punimet mbi to janë përqendruar kryesisht tek modelimi i mënyrës se si tipare të brendshme të përdoruesve mund të merren parasysh nga agjenti. Megjithatë, ka studime që tregojnë se një nga faktorët e rëndësishëm që formëzojnë sjelljen e njerëzve është mjedisi që i rrethon. Edhe pse ka punime që marrin parasysh elementë të mjedisit fizik si për shembull vendndodhja, agjentët e propozuar deri tani përgjithësisht nuk marrin parasysh efektet që mjedisi social ka mbi përdoruesin. Kjo do të ishte e rëndësishme sepse shumica e situatave të përditshmërisë sonë kanë natyrë sociale.

Kjo tezë propozon një strukturë për *social situation awareness*[2] që lejon agjentët e suportit të marrin parasysh situatën sociale të përdoruesit për të ofruar mbështetje gjithëpërfshirëse. Kjo bazohet në përkufizimin tonë të *social situation awareness* në agjentët e suportit: 'aftësia e një agjenti suporti për të perceptuar elementët social të një situate, kuptuar domethënien e tyre, dhe parashikuar efektin e tyre në sjelljen e përdoruesit'. Struktura e propozuar frymëzohet nga studime ekzistuese rreth *situation awareness* në informatikë, si dhe nga koncepte të shkencave sociale.

Fillimisht, propozojmë një arkitekturë konceptuale për *social situation awareness* në agjentët e suportit. Arkitektura bazohet në një set kërkesash që duhet të adresohen në mënyrë që agjentët e suportit të mund të shfaqin *social situation awareness*. Komponenti kryesor i arkitekturës është një modul i përbërë nga 3 nivele: 1) *perceptimi i situatave sociale*, ku perceptohen dhe modelohen veçoritë e mjedisit social, 2) *të kuptuarit e situatave sociale*, ku përcaktohet kuptimi i situatës sociale dhe formohet profili i saj, 3) *projektimi i situatave sociale*, ku vlerësohen pritshmëritë se si situata sociale influencon sjelljen e përdoruesit si dhe vlerat personale. Gjithashtu, arkitektura përmban module bashkëvepruese përgjegjëse për të marrë informacion nga përdoruesit, si dhe për të ofruar suport, shpjegime dhe për t'i lejuar përdoruesve të japin sugjerime për përmirësim.

Pjesa e mbetur e tezës përqëndrohet në implementimin e tre niveleve të *situation awareness*, dhe në mënyrat se si agjentët mund të kalojnë nga njëri nivel tek tjetri. Për të arritur këtë qëllim, kombinojmë koncepte nga shkencat sociale lidhur me marrëdhëniet sociale, karakteristikat psikologjike të situatave dhe sistemeve të vlerave me koncepte teknike nga fusha e *machine learning*[3]. Përgjatë tezës, qasja e propozuar ilustrohet përmes një shembulli hipotetik: një menaxhues kalendarik inteligjent që ndihmon përdoruesit kur ka mbivendosje takimesh. Ky shembull është i përshtatshëm për të ilustruar

[1] Agjentët e suportit janë programe kompiuterike apo aplikacione telefonike inteligjente që ofrojnë mbështetje për përdoruesit, si për shembull Siri në produktet Apple.

[2] Përkthimi më i përafërt në shqip do të ishte 'dijeni rreth situatës sociale'.

[3] Algoritme inteligjentë që përpunojnë të dhënat për të kryer parashikime.

qasjen tonë pasi takimet kanë natyrë sociale, gjë që na lejon të eksplorojmë se si mund të përfshijmë ndikimin e aspekteve sociale të situatave në agjentët e suportit.

Për të realizuar perceptimin e situatave sociale propozojmë që situatat sociale të modelohen nëpërmjet një ontologjie me dy nivele. Ontologjia përqëndrohet në modelimin e marrëdhënieve sociale mes përdoruesit dhe njerëzve të tjerë, si dhe në veçoritë e tjera të situatave sociale. Konceptet e modeluara bazohen në studime në fushën e marrëdhënieve sociale dhe situatave. Në tezë propozojmë një set veçorish të situatave sociale që mund të përdoren për të modeluar takime midis përdoruesit dhe një personi tjetër. Vlefshmëria e këtyre veçorive është provuar nëpërmjet një eksperimenti me përdorues. Për të vlerësuar dobinë e këtyre veçorive kryejmë një eksperiment më të gjerë, në të cilin pjesëmarrësit japin informacion për njerëz nga rrethi i tyre shoqëror nëpërmjet veçorive të propozuara nga ne, dhe më pas u përgjigjen pyetjeve në lidhje me nivelin e prioritetit që do t'i jepnin situatave hipotetike që përfshijnë njerëzit e përshkruar më parë. Me të dhënat e mbledhura ndërtojmë modele machine learning që marrin si input veçoritë e situatave sociale, dhe parashikojnë nivelin e prioritetit të situatave sociale. Rezultatet tregojnë se këto modele arrijnë një saktësi parashikimi që është konsiderueshëm më e lartë sesa modelet bazë.

Më tej fokusohemi se si mund të arrijmë të kuptuarit e situatave sociale në agjentët e suportit. Për këtë propozojmë që karakteristikat psikologjike të situatave, një koncept i studiuar në psikologjinë sociale, mund të shërbejnë si bazë që agjentët e suportit të përcaktojnë kuptimin e situatave. Në veçanti fokusohemi në taksonominë DIAMONDS, që konsiston në karakteristikat psikologjike Ndjenjë detyre (Duty), Intelekt (Intellect), Përballje (Adversity), Partnerizim (Mating), Pozitivitet (Positivity), Negativitet (Negativity), Mashtrim (Deception), Socializim (Sociality). Duke përdorur të dhëna nga një numër i madh takimesh të përshkruar nëpërmjet veçorive të situatave sociale, karakteristikave psikologjike të situatave dhe nivelit të prioritetit, eksplorojmë se si të kuptuarit e situatave sociale mund të implementohet në agjentët e suportit. Fillimisht, demonstrojmë që karakteristikat psikologjike të situatave janë një parashikues më i mirë i prioritetit të situatave sociale sesa veçoritë e situatave sociale. Më tej, eksplorojmë se si mund të përdorim modele machine learning për të parashikuar karakteristikat psikologjike të situatave duke marrë si input veçoritë e situatave sociale. Këto modele parashikuese tregojnë që është e mundur të kalojmë nga njëri nivel i arkitekturës tek tjetri: informacioni i Nivelit 2 mund të parashikohet nga informacioni i Nivelit 1, dhe informacioni i Nivelit 3 mund të parashikohet nga informacioni i Nivelit 2.

Vlerat personale konsiderohen si nxitësit kryesorë të sjelljes njerëzore, dhe për këtë arsye janë thelbësore për agjentët e suportit që duhet të parashikojnë sjelljen e përdoruesit. Vlerat ndikohen nga situatat, prandaj krahas informacionit se cilat vlera janë të rëndësishme për përdoruesin, agjentët duhet të kenë informacion në lidhje me si situatat sociale influencojnë vlerat personale. Konkretisht, në tezë eksplorojmë nëse një situatë sociale promovon, dëmton apo nuk ndikon në aftësinë e një përdoruesi për të përmbushur një vlerë përsonale. Ky informacion do t'u mundësonte agjentëve të suportit të ofronin ndihmë në linjë me vlerat personale të përdoruesit. Vlerësimi se si situatat ndikojnë në vlerat personale është pjesë e Nivelit 3 të arkitekturës - projektimi i situatave sociale. Vlerat personale i modelojmë nëpërmjet një liste të marrë nga studime ekzistuese në shkencat sociale. Më pas, zhvillojmë një eksperiment në të cilin pjesëmarrësit përshkruajnë situata sociale nga përditshmëria e tyre, dhe më pas u përgjigjen pyetjeve në lidhje me karakteristikat psikologjike të këtyre situatave si dhe nëse këto situata promovojnë, dëmtojnë apo nuk ndikojnë në një set vlerash personale. Në të dhënat e mbledhura, grupojmë situatat e ngjashme bazuar në karakteristikat e tyre psikologjike, dhe tregojmë që grupe të caktuara situatash promovojnë ose dëmtojnë në mënyrë të konsiderueshme vlera të caktuara. Për shembull, situatat me një nivel të lartë ndjenje detyre dhe intelekti promovojnë vlera si të qenit i gatshëm për të ndihmuar apo i aftë, ndërsa situatat me nivel të lartë përballje dëmtojnë vlerën e të ndjerit i sigurtë.

Kjo tezë tregon se si komponentët e *social situation awareness* mund të integrohen në agjentët e suportit. Studimet e përmbajtura në të ofrojnë njohuri në lidhje me cilat janë konceptet dhe teknikat që nevojiten për *social situation awareness*, dhe se si ato mund të përdoren në praktikë nëpërmjet shembullit hipotetik të një menaxhuesi kalendarik inteligjent që merr parasysh aspektet sociale të situatave. Ky kontribut është një projekt-skicë për projektuesit e agjentëve të suportit, si dhe shërben si bazë drejt një suporti më gjithëpërfshirës për përdoruesit.

1

Introduction

With artificial intelligent agents being continuously more prevalent in everyday life [84], there is a need for making them better understand and adapt to the users. This involves being able to understand the social environment of the user and its influences on user behavior, a concept which we refer to as *social situation awareness*. The work revolves around the use case of a socially aware agenda management agent, so the focus will be on modelling future situations framed as planned meetings. The aim of this thesis is to improve social situation awareness in support agents. The work stands at the intersection between support agents, situation awareness and sociality (Figure 1.1). In the next sections, we introduce the relevant background knowledge, our research questions, and how they are tackled in the rest of this thesis.

1.1. Motivation and Related Work

Support Agents

Artificial *support agents* that help people in their daily lives, such as personal assistants (e.g., [84]), health coaches (e.g., [160]) or habit formation support agents (e.g., [132]) are increasingly being researched and becoming part of everyday lives, promoted by the increased use of smartphones as well as developments in technologies such as conversational agents. Apart from determining support actions and communicating with users, research on support agents focuses predominantly on modelling personal characteristics of the user, such as their goals, mood, emotions, and personal values [31], [51], [58], [97]. For instance, an online search performed in March 2022 shows that the most downloaded apps that help people quit smoking take into account factors such as cigarette consumption, daily goals, financial and health motivations, but do not consider environmental elements that can lead to smoking. However, research shows that group identity as a non-smoker or adequate support from the social environment are also important contributing factors [112], [113].

The reason for this can be traced back to the work of Lewin [101], who posited that human behavior is a function of the person (including their history, personality, moti-

1

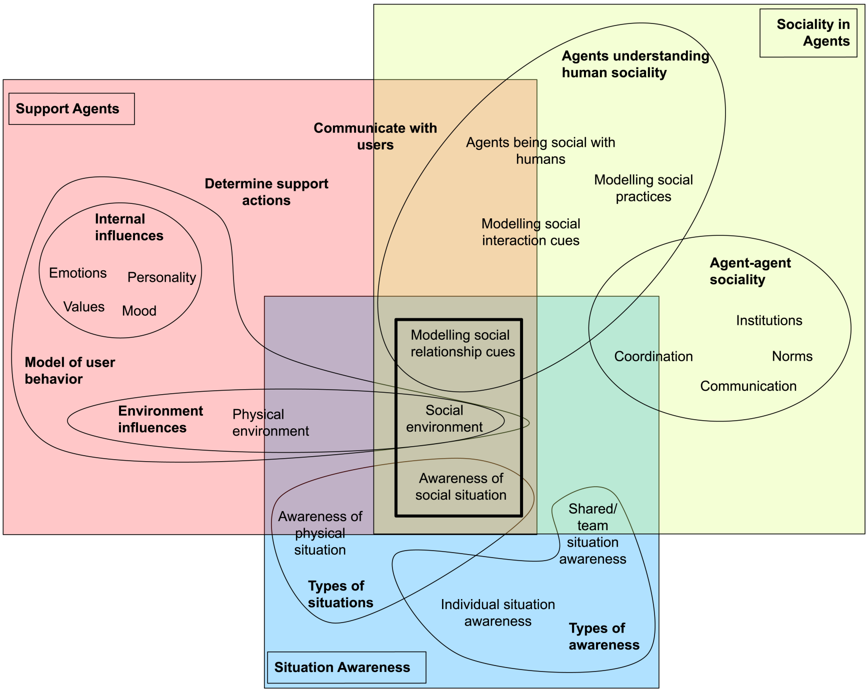

Figure 1.1: Positioning our work in the intersection of support agents, situation awareness and sociality in agents. The thesis focuses on the concepts within the bold square. The list of concepts presented in the image is non-comprehensive.

vation), as well as the environment (consisting of physical and social elements). This suggests that in order for artificial agents to be able to support people in their daily lives, information about their internal aspects is not enough: it is also important to represent the user's current or future situation.

A situation is an interplay of physical and social elements. We focus on the social aspects of situations and how they affect behavior. Humans are social creatures, and many of our daily activities involve other people. Our behavior in these situations is heavily affected by our relationship with the other people in these situations. For instance, dinner with a close friend is in many ways different from dinner with a prospective employer. How to capture and account for the effect of the social elements of a situation on user behavior is an open question in the community. For instance, Van Riemsdijk et al. [166] suggest that support agents should be able to take into account the different social norms from the user's social context. Furthermore, Li et al. [102] identify representing the effect of the social situation on a user as a key challenge for intelligent support technologies. This thesis aims to fill this gap: we propose a framework to enable support agents to take into account how the social situation affects the behavior of the user.

SITUATION AWARENESS

If we turn to social sciences, the concept of *definition of the situation* is considered to be what people use in order to know what is expected of them in a given situation. It is a *subjective* understanding of the role and status of those involved in a situation. We learn how to define situations by combining our experiences with our knowledge of norms, customs, beliefs, and social expectations. The term first appeared in [125], who write: "...In fact, every single act, and eventually all moral life, is dependent upon the definition of the situation. A definition of the situation precedes and limits any possible action, and a redefinition of the situation changes the character of the action."

Research uses terms such as *situation awareness* (e.g. [56]) and context awareness (e.g. [4]) to describe agents' ability to better understand their surrounding environment. According to Barwise [11], these concepts refer to the same thing, and situations represent a way of modelling contexts. Other researchers (e.g. [7]) see context as a lower level of abstraction, and situations can be seen as "logically aggregated pieces of context". In this thesis we focus on modelling a user's situation, since the term is often seen as more comprehensive.

Situation awareness is an over-arching concept used to refer to *'the perception of environmental elements and events with respect to time or space, the comprehension of their meaning, and the projection of their future status'* [46]. This concept helps individuals and teams[1] to process a broad range of situations and facilitates decision making. The most prominent model was proposed by Endsley [47] and consists of three levels: Perception (Level 1), Comprehension (Level 2) and Projection (Level 3). The model has been predominantly applied to understand and describe human decision making in digital technologies that support human situation awareness in domains such as emergency services, military operations, aviation and air-traffic management [156].

Different approaches exist on how to implement the levels of situation awareness in intelligent systems. Situation perception (Level 1) has predominantly revolved around obtaining information from sensors and interpreting that information. The information is usually modelled through domain ontologies (e.g., [15], [104]), and the information from the sensors can be processed using approaches such as fuzzy cognitive maps or Dempster-Shafer theory [15]. Level 2 (comprehension) is predominately represented as situation recognition. For instance, a situation about a user location coming from Level 1 can be used to infer that the user is in a meeting [26]. Different techniques are used to perform this inference step, e.g., rule based approaches such as fuzzy rules [30], and learning approaches such as neuro-fuzzy learning [26]. Level 3 (projection) is inferred from the recognized situation, by determining what is supposed to happen next in the recognized situation. This can be done using techniques such as Gaussian Probability Density Functions [111].

To the best of our knowledge, there are no approaches focusing explicitly on integrating social elements into the three levels of situation awareness. Social situations bring about requirements when it comes to the type of concepts that have to be used and how they can be inferred. For instance, not all information can be easily captured via sensors, since knowledge about social relationships is formed in the user's mind and is not necessarily externalized. Furthermore, situation recognition is not always enough for situ-

[1] This refers to human-human, agent-agent, or human-agent teams.

ation comprehension. For instance, a situation such as 'being in a meeting' can mean different things based on who the user is meeting with. This thesis aims to fill these gaps by providing a framework for situation awareness focused on social situations.

SOCIALITY IN ARTIFICIAL AGENTS

Researchers have argued that AI agents need to incorporate sociality building blocks [81]. The concept of sociality is broad, and so are its applications to artificial agents. In this section, we provide a high level overview of the different ways in which social components have been incorporated in artificial agents. As illustrated in Figure 1.1, the main directions involve agents exhibiting social behavior with other artificial agents, and agents understanding human sociality.

The agent technology research community has explored sociality from the point of view of artificial agents interacting with each other in multi-agent systems governed by structures such as norms, institutions and organizations [37], [52], [107]. In such agent societies, sociality is related to how agents can be organized and how they communicate. Organization structures studied are, e.g., flat structures, where all agents are seen as equal; hierarchies, where agents have tree-like relations; coalitions, where agents are temporarily grouped based on their goals; teams, where agents create a group goal different from their own personal goals [39]. Norms have been suggested as a way to govern agents' behavior through normative systems in supporting coordination, cooperation and decision making [10]. An instance of this is the concept of social commitments, which are norm-based structures that describe agreements between two agents [155]. To enable agent communication, different Agent Communicate Languages have been developed (e.g., [14]). Such languages provide a unique message format and ontology that the agents use to communicate.

The other research direction explores the sociality of agents in relation to humans. This is researched from the perspective of agents modelling human sociality, and the agents interacting socially with people. Modelling human sociality has been explored through the concept of social practices [38]. This involves representing the physical context (resources, places, actors), social context (social interpretations, roles, norms), activities, plan patterns, meaning and competences. The modelled social practices can be used as input to planning techniques in order to determine how an agent should behave in a given social practice. Another approach to modelling human sociality in artificial agents can be found in the social signal processing community and relates to the modelling of social interaction cues such as body language. For instance, Vargas et al. [167] show how social signals measured through sensors can be used as input to machine learning models in order to predict attraction between people. Approaches in which artificial agents socially interact with people incorporates the previously mentioned techniques of understanding human sociality and displaying social signals. A way to explore this has been through embodied social robots, such as social robots that interact with children in the context of education and incorporate psychological concepts such as Regulatory Focus Theory [45] or are capable of social acts such as giving praise [32]. Another way has been through speech or text based conversational agents, such as customer service conversational agents that incorporate natural language processing techniques and social response theory principles [60] or information retrieval agents that use multimodal information to adapt to the user's state and intentions [165].

Less explored is the topic of modelling and understanding the social relationships of humans, such as hierarchy level or quality of relationships, and how to use this understanding for providing support. In this thesis we focus on how to enable support agents to model and understand social relationships of users, and how these concepts affect user behavior and consequently the support agent's decision making about user's social situations.

The CoreSAEP Project

This thesis is part of the CoreSAEP project (Computational Reasoning for Socially Adaptive Electronic Partners), namely support agents that can adapt to diverse and evolving rules of behavior (norms) of people in unforeseen circumstances. Pervasive support technologies while aiming to make our lives more connected and safe, risk violating other values such as freedom or privacy. The project seeks to develop techniques that account for how people's values play out in different situations, and in turn support people in a way that respect their values. The overall vision for the project is introduced in Van Riemsdijk et al. [166]. Part of the project focuses on representing and capturing people's daily activities [127], as well as on a temporal logic to represent daily activities [85]. This provides a basis for representing people's actions in a way that is meaningful to people [16]. This work can be used as a basis for deriving norms that can be integrated in support agents and used to help users, as done in [161] and [86]. This thesis complements these approaches, and focuses on the socially adaptive component of the project. It does so by allowing support agents to also include the influence of the social environment on user behavior, which can then be taken into account in deriving the support actions.

1.2. Research Questions

Our proposed contribution lies in the intersection of support agents, situation awareness and sociality (bolded square of Figure 1.1). This contribution is inspired by Endsley's model [47], which we instantiate with concepts needed to account for social situations. To this end, in this thesis we put forward the concept of *social situation awareness* in support agents, which we define as *"a support agent's ability to perceive the social elements of a situation, to comprehend their meaning, and to infer their effect on the behavior of the user"*. The proposed architecture consists of three levels: social situation perception (Level 1), social situation comprehension (Level 2) and social situation comprehension (Level 3). Specifically, we focus on the elements that a support agent needs to represent in each level, and on the reasoning and information processing techniques that would enable the agent to transition between the different levels, i.e., predict the information of a level by using as input information from previous levels. Our contribution identifies relevant concepts from social science research, models them in knowledge structures, and integrates them in predictive models. This thesis aims to answer the following overarching research question:

> *Which concepts and information processing techniques would enable support agents to exhibit social situation awareness?*

Throughout the chapters we explore the case study of a socially aware agenda management agent which helps its users deal with overlapping meetings. Such an agent would benefit from knowledge on how social aspects affect meetings [158]. This case study was selected because its characteristics allow us to explore the research questions. First of all, meetings that people include in their agenda usually involve other people, making it suitable to studying social situations. Furthermore, the types of possible meetings can be arbitrary rather than about a specific domain, thus allowing us to explore a wide variety of social situations. Secondly, through this case we can study future situations for which the information is available beforehand. This way, we can focus on how the information can be processed to interpret the social situation and its effect on user behavior rather than having to deal with run-time situation perception, since that is beyond the purpose of our current work. Lastly, focusing on an agenda management agent facilitated conducting user studies, since framing social situations as meetings is an easy concept to explain to participants of online studies.

We divided our main research question into five sub-questions. The first sub-question focused on the required high-level components required to tackle the main research question and how they can be combined in a conceptual architecture, and the other sub-questions focused on exploring how each component of the architecture can be implemented. By using research from computer science and social sciences as a basis, we explore which components should be included in a support agent in order to tackle the proposed requirements. We address the following research question:

> ***Research Question 1 (RQ1)****: What components should a support agent architecture include in order to manifest social situation awareness and how can these components be organized in a conceptual architecture?*

Our proposal is a conceptual architecture for social situation awareness consisting of a three level model inspired by the Endsley model of situation awareness: social situation perception, social situation comprehension and social situation projection. The following research questions address how to realize these components.

As a first step towards social situation awareness, the support agent should know which are the relevant elements of the social environment that need to be represented (Level 1 of social situation awareness). Related work by Zavala and colleagues [177] focuses on modelling the concept of places seen as settings where social interactions occur. Other work focuses on modelling social practices [38]. However, these approaches do not capture the full extent of social situations, for instance they do not focus on modelling social relationships. In this thesis, we address the following research question:

> ***Research Question 2 (RQ2)****: Which elements of the user's environment should be modelled in order to represent a present or future social situation of a user?*

In order to fully evaluate the usefulness of the identified elements of the social environment (which we refer to as social situation features), we assess their ability to serve as predictors of user behavior. Specifically, we focus on the priority that the user would assign to different social situations. Priority represents a quantified score of importance and relevance that a user can assign to social situations such as meetings, and it can be used as a tie-breaker when multiple meetings overlap. To perform these predictions, a wide variety of machine learning techniques can be employed with different degrees of success. This is captured by the following research question:

> ***Research Question 3 (RQ3)***: *To what extent can machine learning techniques predict the priority of the situation for the user on the basis of social situation features as input?*

In order to better support the user, the agent should not only know which elements of the situation need to be modeled, but ideally it should also ascribe meaning to each situation (Level 2 of social situation awareness). As mentioned, other work on situation awareness for agents uses Level 2 to perform situation recognition (e.g., [15]). However, we argue that for social situations that is not enough. We propose to ascribe meaning to situations using psychological characteristics of situations, a concept introduced in social psychology [135] according to which people view situations to have characteristics such as duty, intellect etc. We tackle the following research question:

> ***Research Question 4 (RQ4)*** *To what extent can machine learning techniques predict the psychological characteristics of a social situation on the basis of social situation features as input?*

As part of providing human-centered support, it is important for support agents to align its support to the personal values of the user. Aside from determining the priority of social situations, social situation projection (Level 3 of social situation awareness) focuses also on assessing which values are promoted or demoted in a situation. Values are considered a key driver of human behavior [53], and different theories on human values have been proposed (e.g., [143], [150]). Value-aligned support is a desired property of support systems [31], [83]. Our last research question aims to assess how a support agent can determine the promoted or demoted values of a social situation:

> ***Research Question 5 (RQ5)*** *How can the promoted or demoted values of a social situation be determined on the basis of the psychological characteristics of that situation?*

The combination of the insights of the different research questions will be used to draw conclusions about the overarching research question. In the following section we present the thesis outline and explain the approach that was followed to tackle the research questions.

1.3. THESIS OUTLINE

In the rest of this thesis we explore the proposed research questions, and each chapter focuses on one of the above research questions. Figure 1.2 explains the connection between the conceptual architecture, research questions, and thesis chapters.

Chapter 2 explores how the components that are needed for social situation awareness in support agents can be organized in an architecture. An abstraction of the proposed architecture can be seen in Figure 1.2. The conceptual architecture is inspired by the situation awareness model of Endsley [47], instantiated with concepts from social sciences. The architecture consists of three levels, and the remaining chapters focus on how these levels can be implemented.

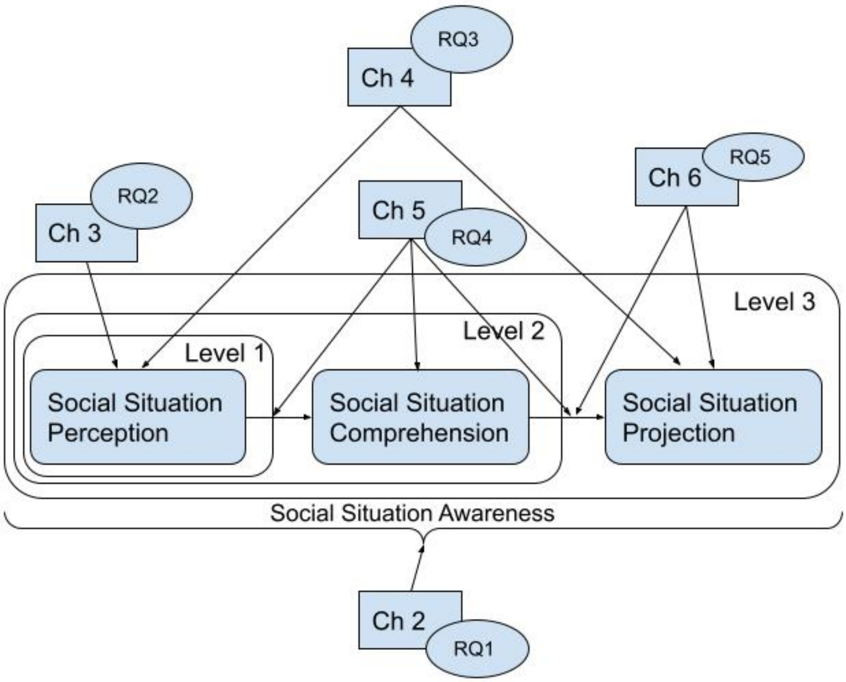

Figure 1.2: Outline of the conceptual architecture and how it relates to the research questions and thesis chapters.

Chapter 3 concerns Level 1 of the social situation awareness architecture, and aims to identify the elements of the user's social environment that need to be represented. To find a solution we explore research from social sciences related to social relationships. Based on the findings, we propose a two-level ontology for modelling social situations. Specifically, the ontology focuses on modelling information about the social relationship the user has with people in the situation (social background features) as well as modelling other social situation cues. The proposed set of features is evaluated through an online user study.

Chapter 4 explores how the proposed set of features can be used to predict the priority of social situations, i.e., predicting Level 3 information from Level 1 information. In the chapter, we conduct an online user study to collect data regarding social situations and their level of priority, and then implement different machine learning models that take as input the features of social situations and predict the priority level.

Chapter 5 explores introducing social situation comprehension (Level 2) as a step in-between social situation perception and social situation projection. We propose using psychological characteristics of situations [135] to ascribe meaning to situations, and explore how the psychological characteristics of a situation can be automatically assessed from Level 1 information. Furthermore, we test whether they can be used as a predictor for the priority of social situations. To explore this, we use data from user studies which collected the features of social situations (Level 1), their psychological characteristics (Level 2), as well as the priority level (Level 3). Lastly, we conduct a qualitative user study to explore how Level 1 and Level 2 information can be used as a basis for explanations given by an agenda management agent.

Chapter 6 explores social situation projection (Level 3). In the chapter we focus on assessing the promoted and demoted personal values of a situation based on the psychological characteristics of the situation (Level 3). To explore this, we collected data from an online study in which people described social situations from their daily lives. Situations were grouped based on their psychological characteristics, and we studied whether groups of situations promote or demote specific personal values.

Conclusions that can be drawn from this thesis are presented in Chapter 7. We discuss the contributions, limitations, as well as societal and ethical impacts of our work. Furthermore, we suggest directions for future work.

2

A Conceptual Architecture for Social Situation Awareness

Artificial agents that support people in their daily activities (e.g., virtual coaches and personal assistants) are increasingly prevalent. Since many daily activities are social in nature, support agents should understand a user's social situation to offer comprehensive support. However, there are no systematic approaches for developing support agents that are social situation aware. We identify key requirements for a support agent to be social situation aware and propose steps to realize those requirements. These steps are presented through a conceptual architecture centered on two key ideas: (1) conceptualizing social situation awareness as an instantiation of 'general' situation awareness, and (2) using situation taxonomies for such instantiation. This enables support agents to represent a user's social situation, comprehend its meaning, and assess its impact on the user's behavior. We discuss empirical results supporting the effectiveness of the proposed approach and illustrate how the architecture can be used in support agents through two use cases.

This chapter was published at IEEE Intelligent Systems [95] (© 2022 IEEE).

2.1. INTRODUCTION

Human behavior is a function of a person's characteristics as well as the situation [101]. Thus, to support people in their daily lives, artificial agents must represent and reason about not only the personal characteristics but also the situation of a user. To take a user's situation into account, support agents should reason about the user's surrounding entities and how they relate to each other and the user.

Most of our daily situations are social in nature. We collaborate with co-workers, spend weekends with family and friends, and share most of our moments with people. Thus, support agents should account for this social dimension of our lives. The importance of social awareness for intelligent agents is also emphasized by recent work in artificial intelligence (AI). For instance, AI researchers argue that agents should incorporate general social intelligence building blocks [81] and account for interpersonal norms [166].

We define social situation awareness and propose the building blocks necessary for support agents to be social situation aware. These building blocks are presented through a conceptual architecture inspired by work on 'general' situation awareness [47], which we instantiate with concepts from social sciences [135] to account for the requirements of modelling social situations. This chapter serves as a proof of concept showing that the building blocks of social situation awareness can be implemented in support agents and discusses the remaining steps for successful deployment of a full-fledged agent.

2.2. WHAT IS SOCIAL SITUATION AWARENESS?

Yang et al. [173] define a situation as "a combination of the individually interpreted, implicit, and unique understandings, and the culturally shared, explicit, and common understandings of the surroundings that produce and constrain human behavior." We define a *social situation* as a type of situation that involves more than one person. Thus, a social situation involves not only the typical situational elements such as time and place but also social elements such as the quality of the relationships and contact frequency between the user and other people in the situation.

The social elements of a situation influence user behavior. For instance, consider two situations: one in which a user has dinner with a friend and another in which the user has dinner with a prospective employer. In these two situations, despite similar environmental elements, the user's behavior can be different because of the different relationships among the people in these situations.

Endsley [47] describes a prominent model of situation awareness consisting of three levels: (1) *perception*, representing the status, attributes and dynamics of relevant elements in the environment; (2) *comprehension*, representing a higher level synthetized meaning of the elements of the environment; and (3) *projection*, representing the ability to project the future status of the elements of the environment. Adapting Endsley's model, we define *social situation awareness* as:

> A support agent's ability to perceive the social elements of a situation, to comprehend their meaning, and to infer their effect on the behavior of the user.

2.3. SITUATION TAXONOMIES

Situations are abstract entities, which makes assigning meaning to them challenging. Studies in social psychology [135] suggest that people interpret situations by creating impressions of them as if they were real entities which have specific *psychological characteristics*. Understanding situations in terms of these characteristics allows people to better navigate the world by using these characteristics to predict what will happen and coordinate behavior accordingly. We propose that support agents should similarly treat situations as real entities with psychological characteristics.

Psychological characteristics provide a high-level subjective interpretation of situations and are widely studied. There are five main taxonomies which provide a set of comprehensive psychological characteristics to describe arbitrary situations [22], [57], [126], [135], [179]. Here we present the elements of the DIAMONDS taxonomy [135]. We choose this taxonomy because it is designed to cover daily situations and it offers a validated scale for measuring the psychological characteristics of situations. The taxonomy comprises the following characteristics in terms of which situations can be described:

- **Duty** - situations where a job has to be done, minor details are important, and rational thinking is called for;
- **Intellect** - situations that afford an opportunity to demonstrate intellectual capacity;
- **Adversity** - situations where you or someone else are (potentially) being criticized, blamed, or under threat;
- **Mating** - situations where potential romantic partners are present, and physical attractiveness is relevant;
- **pOsitivity** - playful and enjoyable situations, which are simple and clear-cut;
- **Negativity** - stressful, frustrating, and anxiety-inducing situations;
- **Deception** - situations where someone might be deceitful. These situations may cause feelings of hostility;
- **Sociality** - situations where social interaction is possible, and close personal relationships are present or have the potential to develop.

Rauthmann et al. [135] suggest that people use these psychological characteristics to ascribe meaning to situations. Furthermore, they show that psychological characteristics of situations correlate with various situation cues, as well as behaviors that people exhibit in those situations. For instance, a high level of Duty is characteristic of work situations, and typical behaviors for situations with a high level of duty are concentrating and displaying ambition. This corresponds to our definition of social situation awareness: psychological characteristics of situations can be used for social situation comprehension, and are related to both social situation perception and social situation comprehension.

2

2.4. Requirements for Social Situation Aware Agents

Different context awareness architectures have been proposed for different purposes. Alegre et al. [6] provide a review of existing approaches, and suggest that one of the reasons for the variety of existing approaches is the need for specific architectures in each domain. However, none of the reviewed approaches tackles social situations specifically. Our research fills this gap. Focusing on social situations motivates us to take into account the human aspects of a situation as opposed to the technical aspects investigated in related work, such as geo-spatial locations and other physical elements of context. Furthermore, the focus of existing approaches is on information that can be acquired through sensors, which is processed to determine actions that are occurring in the environment. Our work complements these approaches by focusing on the psychological meaning of situations. Based on these differences, we formulate the following requirements for support agents to be social situation aware.

1 - Combining sensory data with a user's mental constructs: Perceiving social situations relies not only on information that can be detected through sensors, but also on a user's mental constructs. For instance, in a situation where a user is meeting another person for dinner it is difficult to detect the features of their relationship from sensors alone. This information can be important, e.g., a dinner with a friend is very different from a dinner with a potential employer. Therefore, the agent needs to be able to elicit information about the user's mental constructs such as social relations, which may not be available via sensors.

2 - Variety of social situations: A flexible support agent should be able to represent a wide variety of social situations a user may encounter. To do so, an agent must identify a variety of social dimensions characterizing a situation. Further, the agent should be able to interpret this situation variety by translating social features into abstractions to determine appropriate support, e.g., using pre-specified rules to categorize situations into a limited number of higher-level situations [105] , or using machine learning to derive information that can be used for reasoning about support.

3 - Interpreting the meaning of situations: Existing work on situation awareness addresses the comprehension step by determining how the perceived objects in a situation are interrelated [13] and recognizing the situation type. For instance, if two users are perceived in the same office, the comprehension step would say that the user is in a meeting. However, in social situations, knowing the type of situation is not sufficient to determine the support needed since it is possible to infer different meanings from this information. For example, being in a meeting with a supervisor is different from a meeting with a potential client. Support agents need to be able to distinguish the different meanings of such social situations.

4 - Value-aware support: Agents should provide support that is consistent with the user's goals and preferences. In social science, it has been argued that the essence of a situation is its affordance of human goals and motives [135]. A way to represent human motives are personal values. Values such as independence or success which express what people find important in life have been found to be key drivers of human decisions, and value preferences exhibit cross-situational consistency [150]. Since providing support in social situations is ultimately about aligning with the user's underlying motivations, we suggest the use of values for personalization.

5- Explainability and directability: Support agents need to be able to explain their suggestions to users. For instance, consider an agent that supports healthy lifestyle. If the agent merely suggests the user to avoid going to a party, this advice might be unexpected. However, if the agent informs the user that going to parties usually leads to smoking, which demotes the value of 'health', then the user can make an informed decision. Further, the user should be able to direct the agent on how to act [78]. Continuing our example, the user should be able to inform the agent that since the party is in a non-smoking venue, health would not be demoted. The agent can then use this information in future situations.

Although variants of these requirements are mentioned in existing work, to the best of our knowledge our formulation and approach towards tackling them in an integrated way is new. The key novel elements in our requirements are the consideration of how to ascribe meaning to social situations, the emphasis on user interaction, and a hybrid human-machine approach for social situation awareness and support.

2.5. A Conceptual Architecture for Social Situation Awareness

We identify the core elements, and their interrelations, for creating social situation aware agents by presenting a conceptual architecture (Figure 2.1. The architecture consists of two main components: a social situation awareness component, and a user interaction component. The first is an instantiation of the three-level situation awareness model proposed by Endsley [47] with social concepts. The second comprises interaction modules needed for integrating situation awareness reasoning with the supportive function of the agent. We provide directions for implementing these components.

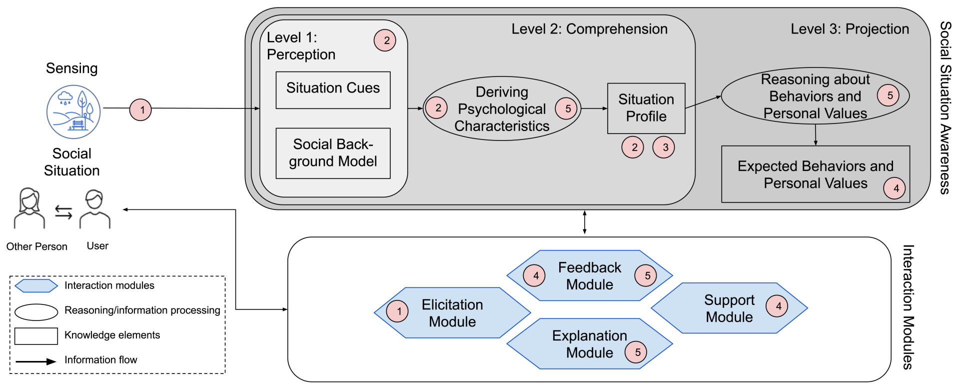

Figure 2.1: Architecture of a social situation aware support agent. The numbers in red circles represent the requirements (Table 2.1) that the elements of the architecture address.

2.5.1. Level 1: Social Situation Perception

The goal of the perception level is to obtain a representation of the salient aspects of a social situation. This information can come from sensory data and interaction with the user. To account for a wide variety of situations, the information included in this level should allow representing arbitrary social situations. In Chapter 3 we propose an approach to model arbitrary social situations through a two-level ontology distinguishing *situation cues* and social relationship features (*social background model*). Rosatelli et al. [144] propose an approach where data from wearable sensors is processed with deep learning techniques to assess information such as roles in social interactions.

2.5.2. Level 2: Social Situation Comprehension

In this level, the perceived elements are used to infer a social situation profile, characterizing the situation along meaningful dimensions.

Knowledge elements

To describe the meaning of a social situation, we propose to use the psychological characteristics of situations (see Section 2.3). The idea is to describe each social situation by a set of features (the *situation profile*) that represent the psychological characteristics of that situation. These characteristics describe a user's subjective understanding of a situation. A key advantage of this approach is that it offers a fixed number of dimensions based on which it is possible to represent and compare different situations.

Reasoning

To determine the psychological characteristics of a situation, one may follow a rule-based or a machine learning approach. A rule-based approach provides explicit reasoning, but requires extensive design time specifications. A machine learning approach supports situation variety, e.g., by offering predictions for unseen examples, but offers limited explainability.

2.5.3. Level 3: Social Situation Projection

In this level, the agent uses the situation profile to predict how the user is likely to behave in a social situation, and what values are affected.

Knowledge elements

In the classic situation awareness model, the projection level captures how the situation develops as a whole. To fulfill the personalization requirement, we propose that in the projection level the agent needs to predict what behavior the user is likely to exhibit, and the personal values promoted or demoted in a given situation. The former allows the agent to provide proactive support, and the latter enables the agent to help the users in a value-aligned manner.

Reasoning

This component takes the situation profiles as input, and predicts the expected behavior and the promoted and demoted values. A possible way to do this is by grouping similar situations based on their profile, and studying the patterns of behaviors and values in each group of situations, as done in Chapter 6.

2.5.4. Interaction Modules

An agent needs to interact with the user to give and acquire information necessary for support. We foresee the need for four interaction modules. In this chapter, we focus on describing the role of these modules as part of the envisaged support agent. In order to realize the interaction modules and create a full-fledged social situation aware agent, open research challenges regarding human-machine hybrid intelligence [3] need to be addressed.

Elicitation Module

The elicitation module interacts with the user to elicit necessary information that cannot be acquired from a sensor. User interaction is needed during both initialization and run time. During initialization, the goal is to gather information that remains relatively stable, e.g., information about a user's social relationships with their most frequent contacts, needed to form the social background model. This ensures that for most social situations which the user encounters, the social background model already contains the needed information, thus avoiding to overload the user with information requests after initialisation. During run time, the module detects when certain information is missing regarding a specific social situation, e.g., the role of the other person, and asks the user. The Platys framework [118] can be used to reduce the possible burden of information elicitation for the user. Platys employs an active learning approach, which asks a user to provide context information only if the predictions with existing sensor readings are uncertain, which reduces the overall effort for the user.

Support Module

After going through the social situation awareness levels, an agent can reason about the support it can provide. One of the proposed requirements is for the agent to personalize support according to the needs and the values of the user. This information can be contained in a user model within the support module. The support module can then compare the user preferences with the information coming from Level 3 of the architecture regarding expected user behavior and values. Support is needed when there is a mismatch between the preferred and the expected behavior of the user in a situation, or when the situation affects a value important to the user.

Explanation Module

To make an agent's support actions explainable, we propose to use meaningful social notions in each level of the architecture, derived through explainable reasoning and learning techniques. An advantage of a multi-level architecture is that explanations can be given on different levels: the agent can (1) give insight on the suggestion relating it to a certain personal value or preferred behavior (Level 3), (2) explain why a certain behavior or personal value is expected in a specific social situation by referring to the psychological characteristics of that situation (Level 2), and (3) give further insight on the situation cues and social relationship aspects that cause the situation to have those specific psychological characteristics (Level 1).

2

Feedback module

It should be possible for the user to notify the agent when a support action or its explanation is not satisfactory. The feedback module achieves this by interacting with the user to determine whether there has been a mistake in one of the reasoning steps or whether some information in the knowledge base needs to be updated. The agent can then integrate this feedback into its reasoning mechanisms and knowledge bases at the appropriate level. How exactly such updates are to be performed and represented is an open research question.

2.6. Empirical Evidence

In this section, we present empirical evidence that supports our proposed three-level social situation awareness component. The social situation awareness component is an instantiation of the well-known model of situation awareness by Endsley [47]. The model's diverse applications in domains such as emergency management [77] and exploratory operations [49] suggest that the three level approach as a whole is beneficial.

Table 2.1: Key requirements and how they are addressed in our proposed approach.

Requirement	**How it is addressed**	**Empirical evidence**
1) Combining sensory data with mental constructs of the user	Perception level based on sensory data and user-elicited information	[92]
2) Variety of situations	Use concepts from social sciences to allow representing arbitrary situations Use machine learning to learn connections between Level 1 and Level 2	[92], [135]
3) Interpreting the meaning of situations	Derive the psychological characteristics of situations	[93], [135]
4) Value-aware support	Base support on the personal values of the user Have feedback module which allows personalization	[94]
5) Explainability and directability	Use explainable techniques Explanation module techniques Feedback module	[93]

In other work, the different levels of the social situation awareness were implemented and evaluated through human-grounded studies [40]. Human-grounded evaluations involve real people who are presented with simplified tasks, and are particularly useful in cases such as ours where the goal is to evaluate reasoning components and a full-fledged agent cannot yet be implemented due to open challenges in interaction modules. In Chapter 5, we show that transitioning through the three levels of the architecture is possible: using data collected from a large user study, we presented an approach in which it is possible to predict Level 2 information from Level 1 inputs, and then in turn use the predicted Level 2 information as input for predicting Level 3 information. Furthermore, we show how Level 1 and Level 2 information can be used as a basis for creating explanations that are satisfying for people. In this section we give details on how the different levels of the social situation awareness module have been successfully implemented and evaluated in generic domains, e.g., to assess the promoted or demoted personal values of a social situation, or specific domains, such as reasoning about the priority of social situations. Furthermore, in Table 2.1 we present how each architectural element tackles the identified requirements.

Level 1

In Chapter 3, we propose an ontology to tackle the perception level. The ontology models situation cues, describing the situation, and social relationship features, describing the relationship of the user with the people in the situation. We evaluate this approach via a user study in which participants were asked about their social relationships using the features proposed in the ontology. Participants considered the ontology to contain an appropriate amount of information (average answer=3, SD=0.61 on a 5 points Likert scale where 1=too little information, 3=appropriate information, 5=too much information) and to be fairly representative of their social relationships (average answer=3, SD=0.79 on a 5 points Likert scale where 1=very little representative and 5=very much representative). This study suggests that it is possible to have a model of a social situation that includes a user's mental constructs, in particular describing social aspects of situations, thus fulfilling Requirement 1.

Level 2

In this level, we suggest ascribing meaning to situations through the psychological characteristics. Rauthmann et al. [135] conducted validation studies involving hundreds of participants across different countries and cultures, showing that the DIAMONDS taxonomy can be used to give meaning to arbitrary situations, thus providing evidence for Requirements 2 and 3. A technical requirement of the architecture is the ability to derive these psychological characteristics from the information from Level 1. To investigate this, we collected Level 1 and 2 data through a crowdsourcing user study (Chapter 5. Using this data, we showed that machine learning models can be created that predict psychological characteristics of situations from Level 1 information with an average error of 1.14 on a 6-point Likert scale, outperforming benchmark results.

Level 3

In Chapter 6, we propose an approach that groups situations based on psychological characteristics and show that different personal values are promoted or demoted in specific groups of situations. For instance, we notice that situations with high intellect and duty promote the values helpfulness and capability. This helps fulfilling Requirement 4. Further, this shows that transition from Level 2 to 3 is feasible with respect to personal values. To show that this transition is also possible in terms of expected behaviors, in Chapter 5 we use psychological characteristics of situations as input to predict expected user behavior regarding social priorities with an error of 0.98 on a 7-point Likert scale for actual values of the characteristics, and with an error of 1.37 for predicted values of psychological characteristics based on Level 1 information.

2.7. Use Cases

We illustrate how the components of our approach could be included in intelligent agents that provide support via two use cases: agenda management [93] and value-based location sharing [83] support agents. Although these use cases are quite different, the high-level components of our approach can be instantiated for each use case as shown in Table 2.2. This illustrates how our approach can serve as a blueprint for including social situation awareness in support agents.

Table 2.2: Concepts that can be modelled and role of modules in two use cases.

Use Case	Agenda Management Support Agent	Value-based Location Sharing Agent
Level 1	Social background features of other person (e.g., role, hierarchy level)	Location-related features; Other people present
Level 2	Psychological characteristics of situation (e.g., Duty, Intellect)	Psychological characteristics of situation (e.g., Sociality)
Level 3	Predict priority of meetings	Assess how values are affected
Elicitation	Social situation features	Personal values
Support	Suggest which meeting to attend based on priority	Provide value-aligned support
Explanation	Why a meeting was suggested	Why a location was shared with someone
Feedback	Adapt priority prediction model	Adapt value assessment

2.7.1. Agenda Management Support Agent

In this thesis we introduce an agenda management support agent, whose goal is to assess a user's priorities and make suggestions based on the priority levels when different meetings overlap. Table 2.2 illustrates the information modelled in the different components of the architecture. Level 1 (*perception*) includes information such as the role of the other person and their hierarchy level. The agent uses the perceived information to assess the psychological characteristics of the situation, which are modeled through the DIAMONDS taxonomy (*comprehension*). From this information, the agent determines a priority level for every social situation (*projection*). If two meetings overlap, the agent suggests to the user to attend the meeting with higher priority and reschedule the other. The user can ask for the reason behind the suggestion, and explanations can be given based on Level 2 or Level 1 information. If the user does not accept the suggestion, the *feedback* module asks about the reasons and incorporates the feedback into the knowledge base and reasoning processes to better predict priority in similar future situations.

In this use case, following our approach allows to explicitly take into account social aspects of the situation, which are modelled from the point of the view of the user through the elicitation module and the perception level. Furthermore, our proposed situation comprehension approach allows for a richer representation and understanding of situations, which in turn allows to better assess priority. For instance an agent may determine that meetings involving a high level of duty are more important for a specific user.

2.7.2. VALUE-BASED LOCATION SHARING AGENT

Kayal et al. [83] propose a model for choosing among conflicting agreements about social sharing of location data based on the users' personal values. They show that an agent can help in resolving conflicting commitments by knowing the value preferences of the user and the promoted values of different location sharing commitments.

Our social situation awareness framework can extend this approach. Level 3 (*projection*) enables the agent to automatically assess which values are promoted or demoted in a situation. Once this information is available, the *support* module can assess whether a specific location sharing activity is aligned with the values of the user. Including information about social relations (*perception*) allows a prediction of values based on a richer model. Furthermore, explicitly modeling the psychological characteristics of the situation (*comprehension*) can be beneficial since these have been shown to be a good predictor of personal values afforded in a situation [94], [135]. For instance, the agent may infer that situations taking place in specific locations involve a high level of sociality, and such situations also tend to promote the value social recognition. If the value is important for the user, the agent would share the location data. This information would also facilitate *explanations*: if the user asks why the location was shared, the agent would explain that it had inferred that the situation promotes social recognition because it involves a high level of sociality. If this inference is not correct, the *feedback* module would adapt the value assessment model accordingly.

2.8. CONCLUSION

In this chapter, we outline the elements needed for social situation awareness in support agents and illustrate their practical benefits. Existing work (e.g., [93], [94], [96]) has shown promising results in implementing the different levels of social situation awareness, as well as in automatically transitioning between the levels using data from studies conducted with real people. This work serves as a proof of concept for social situation awareness in support agents. However, more research from different communities is needed to go from this proof of concept to a full-fledged agent. Firstly, the interactive modules will have a crucial role in the successful implementation of an agent that can be tested on real tasks with users. Realizing these requires further research investigating how hybrid intelligent systems can be made collaborative, adaptive, responsible, and explainable [3]. This includes advances in integrating active learning approaches in order to better personalize the prediction models for specific users based on their feedback. Lastly, this proposed architecture should be integrated with work on interpreting the meaning of social signals such as body language in social interactions [144]. This would allow the agent to take into account the dynamics of a social situation as it unfolds, allowing it to integrate social situation understanding based on social relations with observed social signals.

3

An Ontology for Modelling Social Situations

Behaviour support agents need to be aware of the social environment of the user in order to be able to provide comprehensive support. However, this is a feature that is currently lacking in existing systems. To tackle it, first of all we explore literature from social sciences in order to find which elements of the social environment need to be represented. We structure this knowledge as a two-level ontology that models social situations. We formalize the elements that are needed to model social situations, which consist of different types of meetings between two people. We conduct an experiment to evaluate the lower level of the ontology using feedback from the subjects, and to test whether we can use the data to reason about the priority of different situations. Subjects found our proposed features of social relationships to be understandable and representative. Furthermore, we show these features can be combined in a decision tree to predict the priority of social situations.

This chapter was published in the Proceedings of the International Workshop on Engineering Multi-Agent Systems (EMAS2019) [92].

3.1. Introduction

Artificial agents that support people in their daily lives, for example to live healthier lifestyles or help them in the execution of daily tasks, are becoming a reality (e.g. [120], [160]). Such behaviour support agents need to be aware of a user's *social context* to function effectively [166]: a user's social network may need to play a role in providing support, and a user's activities may involve other people which affects the type of support that is needed [158]. For instance, an app that helps its user be more punctual might send reminders at different intervals when it sees that a meeting is approaching. However, not all meetings have the same priority: for most people, being on time for a job interview is more important than being on time for an informal dinner with friends. Effective support may require taking this into account in the frequency or type of reminders that are generated.

Existing behaviour support agents however mostly focus on modelling internal aspects of the users (e.g. their goals, values, abilities, etc.) [132], [161], while paying less attention to users' social context. In this chapter we take first steps towards developing a *generic framework* that enables behaviour support agents to take into account the user's social environment in order to provide personalized and socially-aware behaviour support [166].

The main idea underlying our approach is to take research on situation awareness, which offers ways to model and reason about the physical environment, and adapt it for realizing *social situation awareness*. Specifically, we take the well-known situation awareness model by Endsley [47] as a starting point. Endsley's model distinguishes three levels of situation awareness: 1) *perception* of relevant elements in the environment, 2) *comprehension* to understand their significance, and 3) *projection* towards future states of the environment. Inspired by these levels, we put forward the idea that a behaviour support agent should similarly be able to represent relevant aspects of social situations, be able to reason about their meaning, and lastly project how these situations will affect the behaviour of the user. These three levels are in line with the classic sense-reason-act cycle in multi-agent systems.

While there are many socially relevant dimensions to behaviour support, in this chapter we focus on handling *social settings* such as meetings or social gatherings. Moreover, we focus on behaviour relevant for arranging these social settings, rather than how to behave whilst participating in one. One may think of a personal assistant agent that can schedule social events for its user [158], or an agent to support people with cognitive impairments in arranging their social life. Furthermore, we need to determine which dimensions of a social situation may be used to interpret their meaning, i.e., what is the "output" of the comprehension process. In this case we focus on *priority* of social situations. We expect that priority, among other things, may be used for dealing with conflicts in a user's schedule. Putting this together, in this chapter we address the following research questions and hypothesis:

- **RQ1** - Which features can be used to describe a social situation from the perspective of a user for the purpose of behaviour support?

- **RQ2** - How can features of a social situation be used to assess its priority?

- **RH** - Priority of social situations can be used for resolving conflicts between two social settings if they cannot both be attended.

While these research questions and hypothesis guide the work presented in this chapter, we do not aim to provide definitive answers here. Rather, as this is a novel research direction, our aim is to assess the feasibility of the approach as a basis for future work that considers other dimensions besides priority, as well as a more extensive investigation into their translation to support actions by the agent.

Addressing these questions involves creating knowledge structures and reasoning techniques for representation and interpretation of social situations, as well as evaluation with users. We further detail this approach and the envisaged software architecture for our support agent in Section 3.2. We present a knowledge structure for describing features of social situations in Section 3.3. We present our user study to evaluate this knowledge structure and gather data for addressing RQ2 and RH in Section 3.4. Our reasoning model for addressing RQ2 is presented in Section 3.5. We conclude the chapter and discuss our findings in Section 3.6.

3.2. Research Approach and Agent Architecture

The overall objective of this work is to assess the feasibility of realizing social situation awareness for behaviour support agents based on the three levels of situation awareness of Endsley [47]. In this way we get a sense of how the different levels of the framework could work together to achieve social situation awareness early on in the research, and it allows us to identify aspects that require a more in-depth study in follow-up research. Specifically, we address the research questions and hypothesis in the following way:

- RQ1: Which features can be used to describe a social situation from the perspective of a user for the purpose of behaviour support?
  - Model building: based on research in social sciences, we propose an ontology for modelling the high-level structure of social situations, as well as a set of low level features that can be used to describe daily life social situations (Section 3.3).
  - Evaluation: Assessing whether the social features identified in the modelling step are suitable, consists of two parts: i) assessing the understandability and expressivity of these features for users; this is important, since we envisage that we will (partly) elicit these features from users through interaction with the support agent, and explaining the support agent's actions to the user requires that these features are meaningful to users (Section 3.4); ii) assessing the usefulness of these features for situation comprehension; this is assessed via RQ2 (Section 3.5).

- RQ2: How can features of a social situation be used to assess its priority?
  - Model building: One may envisage different ways of building a model that can take features of a social situation and derive a corresponding priority,

for example by pre-specified rules, through machine learning, or a combination. Since an important requirement for this model is its explainability for users, in this chapter we choose a learning method that yields an interpretable model: decision trees. To create this decision tree, we collect data from people via a user study (Section 3.4), and then use the data to learn a decision tree that predicts priority of social situations (Section 3.5).

  – Evaluation: We evaluate the predictive capacity of the decision tree by taking a test data set from the data collected for building the model, and evaluating its capacity in predicting the right priority for a specific event based on information about social features of the situation.

- RH: Priority of social situations can be used for resolving conflicts between two social settings if they cannot both be attended.

  – Data collection: First, we ask subjects about their relationship with people in their social circle. Then, we present them with social situations involving these people, and ask them what priority they would assign to these situations. Lastly, we show them pairs of these situations and ask them which one they would attend if the meeting times would overlap and they had to choose only one (Section 3.4).

  – Hypothesis testing: To test this hypothesis, we check whether the proportion of meetings with higher priority that was chosen when breaking the ties is higher than chance.

Figure 3.1 depicts a high level architecture of how our proposed behaviour support agent can be used in practice. The first part of the work consists in learning a model which given data from different social situations (in our case, the experiment data), it learns priority rules based on the answers of the participants. When the user is faced with a future social situation, it gives the behaviour support agent a description of the situation (*situation cues*) and relationship with the other person (*social background features*). The agent uses this information, as well as the learned priority rules, to reason about the priority of this situation. In future work, the priority level will be fed to a support reasoner, which will then output a support action to be of assistance to the user. We hypothesize that priority can be used to break ties when different meetings overlap. In that case, the support reasoner can compare the priority of the different meetings, and suggest to the user which one to attend.

3.3. Modelling Social Situations

In this section we outline which features can be used to describe a social situation from the perspective of a user of the behaviour support agent. We distinguish between a description of the main components, i.e., the overall structure of a social situation (Section 3.3.1) which we refer to as the upper ontology following [72], while the concrete features of the social situation that are the result of the perception process are described in a lower ontology (Section 3.3.2).

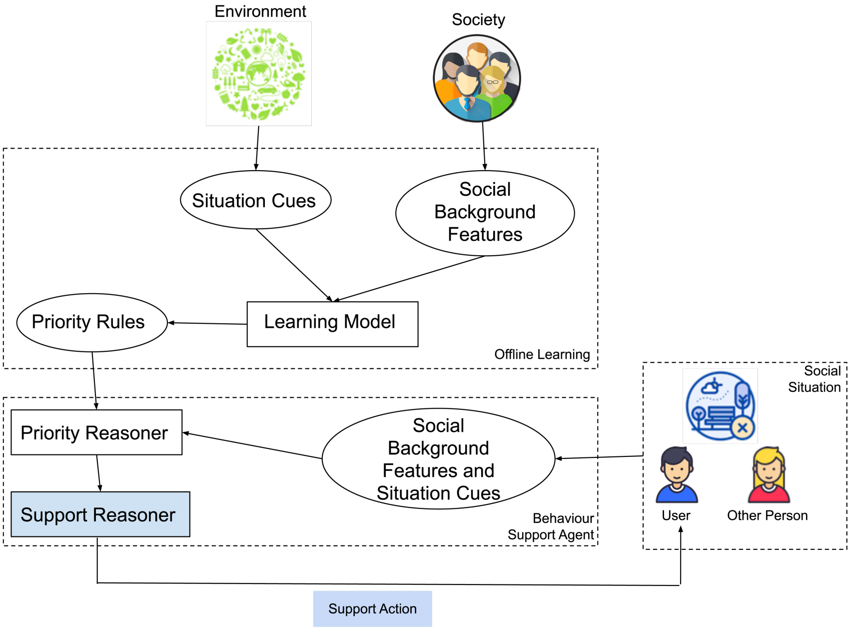

Figure 3.1: High level architecture of the proposed approach. Boxes marked in blue are parts which we do not explicitly tackle in this chapter. Icons used in the image were made by Freepik and retrieved from www.flaticon.com

3.3.1. STRUCTURE OF SOCIAL SITUATIONS: UPPER ONTOLOGY

Research in social psychology by Rauthmann and colleagues [135] proposes that features of situations can be discussed on three different levels: *cues*, which are physical and objective elements (who is present, what activity is taking place, etc.), *psychological characteristics*, which are dimensions that can be used to describe situations (such as duty, intellect, etc.), as well as *classes*, which are abstract types of situations (such as social situations, work situations, etc.). For the scope of this chapter we focus on situation cues and classes, since these are concrete concepts that can be elicited from the user, i.e., that are the result of the perception process. Psychological characteristics, and how to automatically infer them, will be explored in the next chapters.

Cues in turn can be divided into three categories according to Rauthmann et al. [135]: persons, events/activities, and locations. Saucier et al. [149] identify similar categories in an experiment in which students describe their daily situations, namely locations, associations (i.e. people/interactions), as well as actions and positions. Thus we can see that in the literature information about people in the situation is considered to be a specific kind of situation cue. Since in this chapter we focus on modelling *social* situations, meaning that the relation to the people in the situation is of specific interest, we decide to model people separately from other situation cues. This is in line with other work in the field of socially intelligent technologies [1], [2].

3

Cues

The literature identifies essentially two remaining types of cues [135], [149], when we separate information about people from other situation cues: *location* and *activity*. In this chapter we also model the situation class as a type of cue, which we refer to as the *setting*. Furthermore, we introduce a number of additional cues that we consider specifically relevant for comprehension of organized events, as we focus on in this chapter. In particular, we represent the *frequency* with which an event takes place. This variable is not explicitly mentioned as a situation cue in the literature, however some situation taxonomies, e.g., [126], suggest typicality as one of the psychological characteristics of the situation. Moreover, we represent the *time* at which the event takes place, as well as the *initiator* of the event, since we expect this may influence the priority of the meeting.

People

For reasons of simplicity, we focus on dyadic social relationships, i.e., we concern ourselves with social situations involving two people. In our case, one of the people will be the user of the behaviour support agent. This means that the information about the social relation is modelled *from the perspective of the user*.

We model the social relationship by identifying a set of features that characterize this relationship. We distinguish between *social background features* and *situation-specific social features*. The former concern features that describe aspects of the relationship in general, while the latter describe aspects that are specific to the situation at hand. We distinguish two kinds of social background features, namely *structural features* and *personal features*. The former concern what may be referred to as “objective” characteristics such as the user’s role in relation to the other person, while the latter concern “subjective” relationship characteristics from the perspective of the user, such as the quality of the relationship. This distinction is in line with research in social science on relationships in organizations [109] and social support [80], which considers the difference between relationship characteristics that are derived from formal requirements of a role, and interpersonal characteristics. These features are further detailed in Section 3.3.2.

Putting this all together, Figure 3.2 offers a schematic representation of the upper ontology.

Related Work

Context and situations are well studied concepts in computer science. Kokar and colleagues [88] present an ontology for formalization of situations based on the situation theory developed by Barwise [12] and extended by Devlin [33]. This formalization is compatible with the interpretation of situation awareness provided by Endsley [48], which also forms the basis of our work. Yau and Liu [174] offer another ontological approach that models situations for pervasive computing applications. They differentiate between situations, defined as “a set of contexts in the application over a period of time that affects future system behavior" and contexts, defined as “any instantaneous, detectable, and relevant property of the environment, system, or users". Their ontology is based on this division, and they specify a context layer, which models context definition and contextual data, and a situation layer which is built on top of the context layer and aggregates context into situations. This forms the core of their upper ontology, whereas the elements of the lower ontology can be specified depending on the domain.

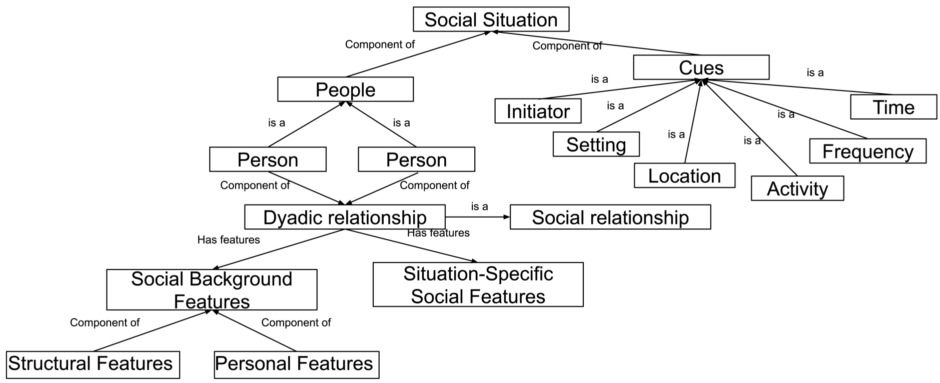

Figure 3.2: Schematic representation of an upper ontology of dyadic social situations.

Their definition of context can be compared to our notion of situation cues. However, these approaches are very abstract in the concepts used in the ontology since they focus on modelling a generic type of situations. Building the lower level ontologies, specifically concerning the modelling of social situations as we focus on in this chapter, is not a trivial task.

Zavala and colleagues [177] offer a framework which can be used to build *place-aware mobile applications*. To do so, they build a place ontology which models the concept of place not only as a geographical location, but also in terms of activities that occur there. For instance, someone can have an office in two different cities, but both of them would count as a *workplace* since similar activities occur there. This is comparable to the cues “location” and “setting” in our ontology. In Murukannaiah et al. [117] this approach is extended and social circles are learned based on the places in which people are meeting: following the previous example, people meeting in workplaces would be classified as colleagues. This can be viewed as a kind of structural relationship feature, as we refer to it in our ontology. Similar to our work, their approach goes beyond modelling very abstract concepts for representing generic situations. However, the concept of places and associated types of relationship is just one aspect relevant to comprehending social situations. Our approach aims at providing a more comprehensive knowledge structure for modelling social situations, as well as development of methods for interpreting these.

Another related line of research is work on modelling and reasoning about *social practices* [35], [38]. In [38] social practices are represented by distinguishing physical context (resources, places, actors), social context (social interpretation, roles, norms), activities, plan patterns, meaning and competences. Physical context and activities are comparable to what we refer to as cues of a social situation, while social context in our case concerns the modelling of people. Meaning can be compared to our second level of situation awareness, i.e., comprehension. Thus while the type of notions we use for modelling social situations are broadly comparable to what is used in research on social practices, our starting point is different. In social practice modelling, the starting point is the social setting, e.g., a classroom [35], for which the norms and expected activities are explicitly modelled *independent of the participating agents*. Then deliberation techniques are needed to allow agents to determine how to achieve their goals, taking into account the (given) norms of this social setting [38]. In our work the starting point is the social relation between the (human) *agents*. For this reason we go in detail regarding the modelling of social background features (Section 3.3.2) that characterize from the (subjective) user's point of view their relation with the other person in the social situation. Based on these features, we then interpret in a bottom-up way the social situation in terms of more abstract general characteristics, in this case priority of a social event. From that we then determine appropriate support actions for the user. Moreover, since our aim is to create behaviour support agents for *people*, we develop our models taking into account results from user studies.

In our previous work [89] we provide an extension of the ontology of [88] with relations that support modelling social relationships, and explore how these can be used for decision making in social situations. However, in that work we model social relations based on only four abstract relationship types from [50] that can be used to model social decision making: communal sharing, authority ranking, equality matching, and market pricing. These can be viewed as a type of structural relationship feature. However, these do not capture personal features that describe more subjective aspects of interpersonal relationships. Moreover, in that work we do not investigate comprehension of a social situation based on these features, but rather model decision making directly using prespecified rules.

3.3.2. Features of Social Situations: Lower Ontology

In this section we go more in detail regarding the modelling of situation cues, and we introduce features of dyadic social relationships that a behaviour support agent can use to model daily life social situations of a user. The list of features presented in this section is not exhaustive, and depending on the type of behaviour support different features may be relevant. However it highlights the type of features that may be considered, and serves as an example of the concrete features that can be used. Moreover, we use these features to model the scenarios in our experiment.

We represent features of social situations by means of relations over situation instances ($\mathcal{SI}$) and dyadic social relationships ($\mathcal{A} \times \mathcal{A}$ where $\mathcal{A}$ is the set of people) for cues and social features respectively, and a domain ($\mathcal{D}$) that specifies the value-ranges the feature can take. This is in line with situation theory ontology [88] in which the modelling of perceived aspects of a situation is done by means of so-called *infons* which describe the relations between objects in a situation. The appropriateness of the chosen value-ranges is also subject to evaluation, and may be changed depending on the domain.

Cues

For simplicity in this chapter we focus on three out of six cues that have been introduced in Section 3.3.1: the initiator of an event, the setting of a social situation, and frequency of the event. A good starting point for modelling locations and activities can be the work of Zavala et al [177].

The initiator is a person from the set $\mathcal{A}$, or `none` if no initiator is identified. For the selection of types of setting of a situation we choose common situation classes that users may face in their daily life. In this chapter, we base the types of settings on Pervin [130], who identifies work situations, family situations, friends/recreation situations, and private recreation situations. We omit the latter since we are concerned with social situations, and add sports activity as a specific type of setting. The situation classes proposed in Rauthmann et al. [135] can also be clustered into these settings. We distinguish two frequencies, `regular` and `occasional`. While more fine-grained distinctions can be made, we expect that this broad categorization suffices in many cases. We list the corresponding relations in Table 3.1 below.

Table 3.1: Relations to model cues of social situations. For a relation $\langle name \rangle$, set of situation instances $\mathcal{SI}$ and domain $\mathcal{D}$, the relation is defined as $\langle name \rangle : \mathcal{SI} \times \mathcal{D}_{\langle name \rangle}$.

Relation name	**Domain** ($\mathcal{D}$)
event_initiator	$\mathcal{A} \cup$ `{none}`
setting	`{work_related, casual_meeting, sports_activity, family_related}`
event_frequency	`{regular, occasional}`

SOCIAL BACKGROUND FEATURES

While there is a lot work in the social sciences on understanding social relationships, in this chapter we mainly use the following two lines of work as the basis for selecting structural and personal social features for our model. First, Kahn and Antonucci [8], [9], [80] explore the role of social relations as a form of social support for (elderly) people. Enabling social support is an important purpose of the behaviour support agents we aim to create [166]. We select our structural features mainly from this line of work. Second, social relations are also considered from the organizational point of view. Specifically, we use the work of Mainela [109] which gives an overview of types and functions of social relationships that can be relevant in the organization of a joint venture. Organizational relationships are an important type of relation that our behaviour support agent may take into consideration. We select our personal features mainly from this work.

Structural Features Kahn and Antonucci conceptualize support systems as a so-called Convoy model - three concentric circles representing three levels of closeness between the supported person and their "convoy" of supporters. Different aspects of the relationship are considered in order to establish someone's position within the convoy model. The Convoy model [9] distinguishes between structural (age, sex, years known, proximity, contact frequency, relationship (role)) and functional characteristics (types of support received and provided) of social support networks.

For this chapter we use *role*, *contact frequency*, and *default geographical distance* (proximity) as structural features. The feature *role* refers to the role of the other person towards the user in dyadic relations. Knowing this is important since it can help inferring the expectations that come with the role. The range of roles we use is taken from the general social survey [23]. The geographical distance refers to the physical proximity of the two actors in terms of their default home location. Proximity can influence the relationship of two people since it affects how often they can see each other. For the range we opted to measure distance in terms of time that it usually takes to get to that person.

Besides the above three structural features, we introduce a fourth one, namely *hierarchy*, to express the type of relation between the user and the other person. Hierarchy affects the power dynamics between the first and second actor. Higher (respectively same and lower) means that the other person is higher up (resp. at the same level, and lower) in the hierarchy than the user. In case there is no hierarchy amongst the actors, this is indicated by “n.a.". We expect this feature to be relevant when assessing the priority of meetings, especially for users who are in working relations, or actors that come from a culture with some sort of caste system. More information on the concept of hierarchical ranking can be found in, e.g., [50], [171].

Personal Features The first of our personal features is also taken from the Convoy model [8]. In addition to structural and functional aspects of relationships, they emphasize the importance of *relationship quality* in characterizing social relations.

The remaining three personal features we consider in this chapter are taken from Mainela [109]. They give an overview of how types of social relationships in business dyads have been characterized in the literature. For example, Granovetter [63], [64] talks about strong ties and weak ties in work relationships. The strength of a tie in a network depends on aspects such as the amount of time spent on it, the emotional intensity, the intimacy, and the reciprocity. Furthermore, the author argues that ties are stronger when the level of acquaintance is deeper.

From the list of features for characterizing social relations identified through the literature study of Mainela, we select three, namely *acquaintance depth* [63] of the user towards the other person, *level of formality* of the relationship [140], and *trust* [164] of the user towards the other person as personal features. These features can inform the expectations of the relationship between user and the other person, and consequently are relevant for comprehending social situations.

Other features mentioned by Mainela can be used to distinguish different types of social relationships in a business context, but seem too specific for social situation awareness of our envisaged behaviour support agent, e.g., legal questions, attendant consequences, activation of a relation, outcome expectations, and scope of economic issues. The features continuity of interaction and amount of time spent are closely related to event frequency, contact frequency and acquaintance depth. Features like personal nature, intimacy and emotional intensity seem closely related to level of formality and acquaintance depth. Finally, reciprocity may also be relevant for our purposes, however refers more to functional aspects of the relationship and may be difficult to characterize directly in these terms by users. Therefore we leave it out in this chapter.

We summarize these social background features in Table 3.2 below. The range of some features is $\texttt{Likert}_5$, which denotes a 5-point Likert-type scale, where 1 is the lowest/most negative value and 5 the highest/most positive value.

Table 3.2: Relations to model social background features of social situations. The upper part concerns structural features, the bottom part personal features. For a relation $\langle name \rangle$, and domain $\mathcal{D}$, the relation is defined as $\langle name \rangle : \mathcal{A} \times \mathcal{A} \times \mathcal{D}_{\langle name \rangle}$ where $\mathcal{A}$ denotes the set of persons.

Relation name	**Domain** ($\mathcal{D}$)
role	{partner, parent, sibling, child, extended_family, coworker, neighbor, friend, supervisor group_member, other}
contact_frequency	$\texttt{Likert}_5$
def_geo_distance	{0-1hr, 1-2hr, 2-4hr, flight needed}
hierarchy	{higher, same, lower, n.a.}
rel_quality	$\texttt{Likert}_5$
acq_depth	$\texttt{Likert}_5$
rel_formality	$\texttt{Likert}_5$
trust	$\texttt{Likert}_5$

SITUATION-SPECIFIC SOCIAL FEATURES

Several of the social background features may have a situation-specific variant, for example if you go to a basketball game with your boss, in that situation you are both teammates, and if you are the captain you are the one holding a higher hierarchy level in that situation. However for reasons of simplicity we do not further elaborate on these in this chapter.

We do introduce another situation-specific social feature, which we call the *help dynamic*. It refers to whether in the specific event the user is giving to or receiving help from the other person. The fact that they have to give or receive help can influence how obligated the actors feel to attend a certain event. It is defined as a relation *help_dynam*: $\mathcal{SI} \times \mathcal{A} \times \mathcal{A} \times \mathcal{D}_{help\_dynam}$, where $\mathcal{D}_{help\_dynam} = \{\texttt{giving, receiving, neither}\}$.

RELATED WORK

Different aspects of modelling social relationships have been studied in sub-fields of multi-agent systems. In particular, when talking about organizations of agents, "role" is one of the central concepts. In the OperA model [37], agents form societies with different organizational structures, and they take up roles in these societies. These roles, in combination with social contracts, define what an agent should and should not do. Singh [155] follows a similar approach, and proposes that "Org(anization)s are finely structured through the notion of a role, which codifies a set of related interactions that a member of an Org may enact". D'Inverno and colleagues [43], in their quest to weave a fabric for socially aware agents, also introduce the concept of roles in order to represent agents in the context of a social setting. Roles in these works are used to describe, design and understand interactions in an abstract and re-usable sense, independent from the agents that will eventually play the roles. In our case we combine abstract information about roles with information about the concrete relation between the user and the other person, i.e., between the specific (human) agents in the interaction, in order to assess how best to support the user this social situation.

The notion of hierarchy is used in [37] to describe a type of relation between roles in an organization. Although not the same thing, hierarchy can be connected to the notion of power. Pereira and colleagues [129] argue for the importance of modelling social power into the decision making of cognitive agents. The importance of modelling social power is also proposed in [36].

Another well studied concept within the multi-agent systems field is trust. Mostly, it is considered from the point of view of software agents trusting each other. The focus is on determining the level of trust in another agent by taking into consideration the agent's previous interactions with another agent, or by relying on other agents' opinions about that agent [55], [176]. In our case, once we have information about the trust the user has towards the other person, we use it for interpreting the social situation and allowing our support agent to determine the appropriate support actions in this situation.

The virtual agents research area has also studied modelling and use of various features that describe social relationships. Zhao and colleagues [178] argue for the importance of representing rapport in a virtual agent that interacts with a human. Rapport is a feeling of connection and closeness to another person, which can be compared with depth of acquaintance. Dudzik and colleagues [42] provide a review of literature that deals with contextual features of human emotion perception for automatic affect recognition. As contextual factors they identify characteristics of the sender or receiver of the emotion, such as age, gender and occupation, as well as situation features such as cause of the emotion, conversation content and language, information about the conversation partner in the social interaction, location, and lighting conditions during the interaction. Our work is complementary in that it focuses on characterizing the social relationship itself between people in the social situation, and from that derive higher-level understanding of the social situation, in this case in terms of its priority.

Thus our framework for modelling social situations includes a number of features that have been studied in various parts of the agent systems literature. Based on social science literature we add several features that are specifically relevant for characterizing human social relations, such as contact frequency, geographical distance, and relationship formality. Moreover, our work differs from existing work in multi-agent systems in that we investigate how we can *combine* features of social situations for the purpose of comprehension in order to allow an agent to provide appropriate socially-aware support.

3.4. User Study

In order to evaluate how well we can use our proposed low level features to model and interpret daily social situations, we conducted a pilot experiment in which subjects had to answer a survey about the social relations in their life [90]. The survey consisted of three parts, through which we explore RQ1 and evaluate RH. Furthermore, we use the data from survey to create and evaluate a model that addresses RQ2. We present our experimental setting in Section 3.4.1 and our results in Section 3.4.2.

3.4.1. Experimental Setting

Pilot Subjects

We tested 20 subjects (15 male, 5 female) who answered to all three parts of the experiment. Subjects were university employees (mostly PhD candidates). The average age was 31.1 years old (SD=7.6yo).

Design and Procedure

The experiment[1] was implemented as an online survey, and consisted of three parts.

In *Part I*, subjects were asked to think about six people from their social circle. For the purpose of the study, they were instructed to select at least one family member, one friend, and one person who had a higher hierarchy level than them. In follow-up research, we will also ask for information on relationships with people lower in the hierarchy. For each of these people, subjects were asked to provide all social background features (Section 3.3.2). The first part was concluded with an evaluation section in which the subjects were asked whether the questions were understandable, whether the amount of questions was appropriate, and how well they thought the questions represent their social relationship with someone. Through these questions, we test how understandable and expressive our proposed features are (RQ1). Furthermore, they had the option to propose more aspects of social relationships which they thought are relevant.

In *Part II*, subjects were shown 20 scenarios of daily life social situations. Each scenario involved one of the six people that subjects had mentioned in Part I, selected randomly[2]. We made the study subject-specific to enable them to reflect on their own relationships, instead of presenting them with hypothetical relationships. Scenarios consisted of different parameters of the situation cues and situation specific features of social relationships. A scenario could represent a social situation such as:

> “You have invited *Person X* for a work meeting on Tuesday morning because you need some feedback on your recent project".

In this case it is a work setting, the event is occasional, the subject is the initiator and he/she is expected to receive help. For each scenario, subjects were asked about the priority of the meeting, how obligated they would feel to attend the meeting and how much they would enjoy it. We need the information on priority to answer RQ2. Obligation and enjoyment were asked for exploratory purposes to inform future research. Furthermore, subjects were asked how they think the other person would answer these questions. This was done because in future work, we want to explore the reciprocity of these decisions. Lastly, they were asked about the likelihood of that scenario happening in their daily life in order to assess the appropriateness of the scenarios we have chosen. Subjects had to answer on a 5-point Likert scale. In order to assess priority, they were instructed to take into account how difficult it would be for them to cancel the meeting, how important they think it is to be punctual, and any other thing they would consider relevant.

In *Part III*, scenarios were paired randomly and subjects were asked which of the two meetings would they choose to attend in case of a conflict between the two scenarios meaning that they could not attend both meetings. We will use this information to evaluate RH. Furthermore, they were asked what reason would they give to the person whose meeting they were canceling: the real reason, some other reason, or no reason. This was

[1] The questions for each part of the experiment can be found in the Appendix. The data can be accessed in https://doi.org/10.4121/uuid:e18fb318-c1d4-4ccc-9b4f-be48e1ee49e2

[2] Apart from the scenarios in which a family setting or a higher hierarchy work setting were being tested, which were restricted to family members and people with higher hierarchy, respectively.

asked in order to have some more insight in case our hypothesis is not corroborated from the data. Each subject was presented with six pairs of scenarios.

3.4.2. RESULTS

In this subsection, we present and discuss the results of each part of the experiment separately.

PART I

The selected people from the subjects' social circle had an average age of 37.6 years old (SD=13.55yo). They were mostly friends (29%), followed by people from work (18% supervisors and 10% coworkers) and family members (11% parents, 8% siblings and 7% members of the extended family). Partners consisted of 10% of the selected people. Overall 74% of the people were not in a hierarchical relation with the subjects, 22% were on a higher level and 4% on a lower level. 36% lived within an hour of distance from the subjects, 18% between 1-2 hours, 4% between 2-4 hours, and for the remaining 32%, the subjects would need to take a flight in order to meet them. The subjects' answers for social background features that have a Likert-scale as the domain are shown in Table 3.3 below.

Table 3.3: Percentage of subjects that gave each specific answer for different social background features. The answer options were Likert-type scale values ranging from 1 to 5. For relationship quality 1=very negative and 5=very positive. For the rest, 1=very low and 5=very high.

Feature\Answer	1	2	3	4	5
Contact frequency	0	23.82	25.59	25.59	25
Relationship quality	2.06	5	10.88	49.41	32.65
Acquaintance depth	0	15.59	34.12	25.59	24.71
Relationship formality	46.76	16.18	26.18	7.06	3.82
Trust	1.76	1.47	25	37.25	34.41

As seen in Table 3.3, subjects mostly choose people with whom they have strongly positive relationships. Furthermore, they chose people whom they trust, and the relationships have a low level of formality. In future work, in order to have more representative data from a larger variety of relationships, we will control some features when asking the subjects to think of people from their social circle. For instance, we will ask some subjects to think about a coworker with whom they do not have a positive relationship.

The evaluation questions (all posed with a 5-point Likert scale in possible answers) showed that the subjects found the questionnaire understandable, with an average of 4.59 (SD=0.51). The number of asked questions was appropriate (the average answer was 3, SD=0.61, on a 5-point scale where 3=`appropriate`). When asked how much this information represents their relationship with someone (Likert range from 1 = `very little` to 5 = `very much`), the average answer was 3 (SD=0.79), confirming that social relationships have subtle aspects not captured in our questionnaire. Whether we need to add more features, depends on the strength of the correlations between the current features and the choices the subjects make in Part II of the questionnaire. The subjects (mostly being PhD students), seemed to understand this point, as some subjects indicated that the answer to this question depends on the purpose of the study. This is something that we will take into account in future experiments.

When asked whether they could think of additional aspects of social relationships which should be present in the survey, 35% of subjects answered with "Yes". Some of the suggestions included: dependability, understanding, fun, respect, how important is the other person, common interests, etc. However, none of the suggestions appeared consistently.

PART II

In this section subjects were asked to evaluate different scenarios with respect to their priority (and additionally obligation and enjoyment). Subjects mostly give a high priority to the meetings, with 37% of scenarios being assigned a 5, 41% a 4 and 16% a 3, with only 6% having a 1 or a 2. This was expected given that scenarios included people with whom the subjects have a close and positive relationship. This is also reflected in how much they enjoy these meetings (65% of scenarios being assigned a 4 or a 5). For obligation, the results were more balanced, with 14% of scenarios being assigned a 2, 21% a 3, 37% a 4 and 25% a 5. The average likelihood of the scenarios was 3.14 (SD=1.42), which means the scenarios were relatively likely despite being chosen randomly in terms of the combination of person with whom the subject relates, and scenario. We notice a high standard deviation, caused by the fact that some of the scenarios had a low likelihood, possibly because of the random person-meeting combination.

PART III

In this part, subjects were given pairs of scenarios (from Part II), and they had to select which one they would attend if they could attend only one. We notice that in 69% of the cases, subjects would select the meeting to which they had assigned a higher priority in Part II. This suggests that priority is a good indicator of how people break ties. However, it is not the only thing. We noticed that in most of the cases in which subjects select meetings to which they had assigned a lower priority, those meetings have also a low likelihood. This suggests that when breaking ties between different meetings, subjects also take into account how difficult it would be to reschedule each of the meetings. Also, in this section we see differences between individuals, since there were subjects who consistently chose a certain type of meetings. This can link to the subjects' *personal values* (see also [83], [161]). For instance, some subjects consistently picked work meetings or family meetings, which indicates a tie to their value system. This will be explored in future work.

Subjects were also asked about the justification that they would give to the person whose meeting they would cancel. In 89% of the cases, subjects reported that they would give the real reason. Most of the cases in which the subjects would give no reason or a different reason (and not the real one) took place when they chose to attend meetings with a lower priority. Furthermore, many cases involve either not reporting to someone with a higher rank, or not giving details about their meetings with family members.

3.5. Predicting Priority of Social Situations

In order to address RQ2, we investigate how to use data from Part II of the user study in order to predict the priority level of social situations based on information about social features. First we discuss possible options on how to achieve this (Section 3.5.1), and then we introduce and evaluate our proposed approach (Section 3.5.2).

3.5.1. Reasoning about Situations

Different strategies can be used to reason about the priority of an event. The most straightforward approach would be to combine the situation cues in an Expected Priority (EP) function, such as:

$$EP = \sum_{f \in \mathcal{F}} w_f v_f$$

where $\mathcal{F}$ is the set of all features considered, and where for all $f \in \mathcal{F}$, v_f refers to the feature value and w_f to the relative weight of feature f in this computation. However, there are two main issues with this approach. First of all, most of the features that we are dealing with have nominal values, so quantifying them is difficult. Furthermore, based on the literature on preference profiles, see e.g., [18], in many decision situations, we hypothesize the weights to be dependent on the individual, making the correct initialization of the weights a challenge.

Another option is to learn a model from our data, and use it to classify new instances. Our proposed approach to do this is to use decision trees [21], because literature suggests that the structure of decision trees is appropriate for reasoning about social relations. First of all, cognitive psychology proposes that social intelligence can have a modular nature [59]. This means different "scripts" are activated in different settings. People recognize these settings from environmental cues, and in turn decide to behave in a certain way. This is similar to the concept of decision trees, in which different combinations of features lead to different decisions. Endsley also suggests that people use different "schemata" to organize and combine knowledge and perceptions in order to comprehend the situation [47]. Moreover, the decision process of decision trees is predictable and transparent. This would allow the agent to *explain* to the user why a certain priority level is assigned to a specific event, which is important since we focus on behaviour support.

Decision trees are graphical representations of a set of rules which can be used to make classifications. Each node of the tree represents a question regarding certain features of the object that is being classified, in this case a social situation, and each branch represents a different answer to that question, in this case the priority level. Nodes below a given node either contain another question, or are given a *label* which assigns a class to the object. The latter are called *leaf-nodes*. Given an object with a set of features and a decision tree, in order to classify the object we traverse the tree until we reach a leaf.

3.5.2. Model

So far, we have represented the features of the social situations. However, this raw information is not sufficient to draw conclusions about how people evaluate situations. As explained in Section 3.1, in situation awareness literature, this process is called *compre-*

hension [48]. In this work we explore one general and abstract characteristic of a given situation, namely its *priority*.

As mentioned in the previous section, we use decision trees to predict priority of social situations. One of the most used methods because of its high accuracy is the Classification and Regression Trees algorithm (*CART*) [21]. CART models are binary trees, which means for every parent node there are two child nodes. Learning a CART model involves selecting features and split points on those features until a suitable tree is constructed. This selection is performed by using a greedy algorithm which minimizes a cost function. We build the model using the R package `rpart` [159]. We use 70% of the data as a training set from which the tree structure was learned, and then test it on the remaining 30%. As a pruning mechanism we limit the maximal depth of the tree to 4.[3]

3

The learned model is shown in Figure 3.3. We remark that, to us, many of the tree splits are intuitive. For instance, the first information that is checked is the setting of the meeting, with casual and sport events on one hand (the left branch) and family and work events on the other (the right branch). This split was to be expected since subjects assigned higher priorities to family and work events.

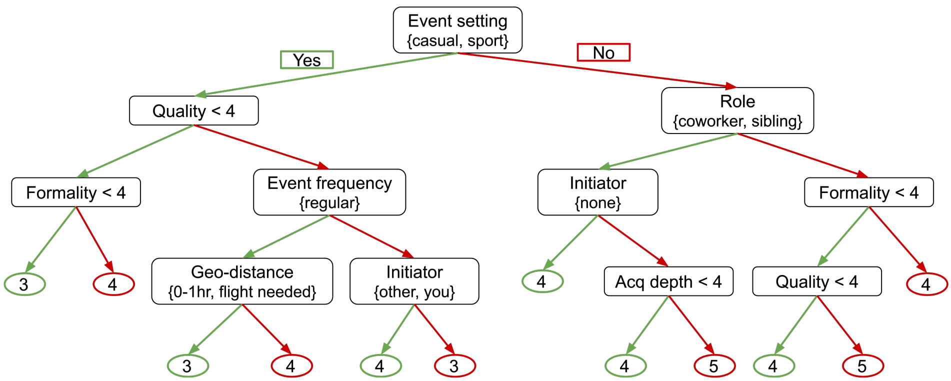

Figure 3.3: Decision tree built based on the data. Nodes with categorical features, such as event setting, should be interpreted as “is *event_setting*=`casual OR sport`?"

Since we lack a benchmark in this domain in order to evaluate our model, we compare our result with an algorithm which would predict a random priority (as we offered 5-point scale, chance corresponds to 20%) and with an algorithm which always picks the most selected class, i.e., priority 4, which was selected in 41% of cases. To determine the accuracy of the models, we use the following definition:

$$accuracy = \frac{Number\_correct\_predictions}{Number\_overall\_predictions}$$

The accuracy of our model on the test set is 47%, thus performing better than the other two algorithms that we used as a benchmark. This means that information about social features can be used to predict priority of a social situation.

[3]The code can be found in: https://github.com/ilir-kola/decisiontree-socialsit.git

3.6. Conclusion

3.6.1. Research Questions and Hypothesis

Regarding RQ1, in Section 3.4.2 we notice that subjects find our proposed set of features understandable and their quantity appropriate. Furthermore, they find the features relatively expressive. In Section 3.5, we tackle RQ2 by proposing a model which learns a decision tree to predict priority of meetings. We observe that the model performs better than chance, which shows that while this can be a way to predict priority of social situations, more works needs to be done in order to achieve a higher accuracy. This may involve introduction of additional features. The result also contributes in the answer of RQ1, since it suggests that the features allow us to represent social situations in order to learn information about them. Regarding RH, in Section 3.4.2 we see that in 69% of the cases, priority is a good predictor for choosing between overlapping meetings. However, it also shows that it is not the only element, and more dimensions of situations need to be assessed to identify where this difference comes from.

3.6.2. Contributions

For the benefit of the development of behaviour support agents with social situation awareness, this chapter provides the following contributions:

- an upper ontology for representing the salient situation cues and types of features for characterizing dyadic social relationships.
- a set of lower level features which can be used to represent daily life social situations.
- an evaluation of social features via a user study, showing that subjects find the concepts understandable and expressive.
- an evaluation whether decision trees can be used to predict the priority of social situations based on features of social situations, which proved to be the case.

Results presented in this chapter tend to support the feasibility of our overall approach, but in parallel they open the way for different research questions which need to be explored in more depth in future work.

3.6.3. Limitations

First of all, the number of people in our user study via which we evaluate our proposed features is relatively small, and the subjects are mostly PhD candidates. This does not allow for a conclusive answer when it comes to understandability and expressiveness among other types of people with, for example, other levels of education. In turn, this also creates limitations when tackling RQ2. First of all, we built the model using a small data set, and learning algorithms need more data in order to generalize better. This is also shown by the high level of over-fitting which takes place, as noticed by the fact that the accuracy on the training set is 65%. Moreover, the data is unbalanced, since people mostly give a priority of 4 or 5 to events. The presence of lower priorities would make the evaluation of the algorithm more realistic since we would be able to measure not only the number of correct predictions, but also how far off the incorrect predictions are. The low variance in the data can be explained by the fact that subjects chose people who are very close to them, thus they would prioritize those events.

3.6.4. Proposed Future Work

Based on the findings reported in this chapter, a more extensive experiment can be confidently carried out to obtain a detailed social model that can serve as a background model for behaviour support agents to advise on how to choose between social situations. More data can help not only in building a more accurate model, but also to try out more techniques. Furthermore, that data can also be used to study the correlations between the different features, in order to select a minimal set of features for which to ask the users.

Another interesting approach is to analyze how *personal values* [83], [161] affect the way in which subjects think about social situations. Part III of our experiment suggested the existence of individual differences in how people decide which meetings to attend. In the next chapter we explore whether people with shared personal values make similar choices.

The current model relies fully on information that is acquired directly from the users. In future work, we would like to add sensory data to inform our model. Literature shows that sensory data can be used to perceive social information (e.g., [24]). This line of research would provide useful ways to acquire information without interrupting the user.

Finally, in this chapter we mostly focus on the modelling of social situations. The next step is to dive deeper into situation comprehension, and reason about different dimensions of social situations (other than priority). Data from the user study suggests that both enjoyment and obligation correlate well with priority, and this correlation is stronger when considering situations in specific settings (enjoyment for casual situations and obligation for work situations). Representing more dimensions of social situations would lead to having a more complete profile of the situation, which in turn enables behaviour support agents to provide more comprehensive help.

4

Predicting the Priority of Social Situations

Personal assistant agents have been developed to help people in their daily lives with tasks such as agenda management. In order to provide better support, they should not only model the user's internal aspects, but also their social situation. Current research on social context tackles this by modelling the social aspects of a situation from an objective perspective. In our approach, we model these social aspects of the situation from the user's subjective perspective. We do so by using concepts from social science, and in turn apply machine learning techniques to predict the priority that the user would assign to these situations. Furthermore, we show that using these techniques allows agents to determine which features influenced these predictions. Results based on a crowd-sourcing user study suggest that our proposed model would enable personal assistant agents to differentiate between situations with high and low priority. We believe this to be a first step towards agents that better understand the user's social situation, and adapt their support accordingly.

This chapter was published in the Proceedings of the International Conference on Principles and Practice of Multi-Agent Systems (PRIMA2020) [96].

4.1. INTRODUCTION

Artificial agents that play the role of personal assistants are increasingly becoming part of everyday life (e.g. [84]). These agents have focused on representing internal aspects of the user, such as their values, goals, or emotions [132]. However, research in social science suggests that human behaviour is shaped both by these internal aspects, as well as by the situation someone is in [101]. Situations have a physical aspect (e.g., where it takes place) and a social one (e.g., who is involved). We focus on the latter: our goal is to build methods which allow personal assistant agents to model the social situation of a user, and use that information to reason about how to provide socially-aware support.

The need for enabling intelligent support agents (such as personal assistants) to understand the social situation of the user has been acknowledged as one of the main open questions in agent research [81], [166]. Existing work on modelling social context focuses on modelling the social practices of a situation (e.g. [35]), or the place where the interaction is taking place (e.g. [117]). In our approach, we model situations from the perspective of the user of the personal assistant agent by modelling how the user relates to the people in that situation on a number of relevant dimensions. This complements [35], which models the social practices *of a situation*, while we focus on modelling the perspective of an individual *on that situation*. This requires additional social features to describe social relations between people that go beyond their roles in the situation. Based on information about how the user relates to the social situation, we investigate how an agent can interpret that situation in order to determine desired actions that can support the user. Regarding our technical approach, we combine the strengths of existing work: we propose an explicit model of a social situation (similar to [35]), and combine it with learning techniques to derive new information (similar to [117]).

To illustrate our approach, we take the example of a personal assistant agent which helps busy users manage their agenda automatically (e.g. [119]). We consider each meeting to be a social situation. The agent takes as input situation cues (e.g. the setting of the meeting, such as a work meeting) and relationship features (e.g. the quality of the relationship, such as a very positive relationship). Based on this the agent determines which meeting the user would likely want to attend when two meetings overlap. If the user is too busy to respond to meeting requests themselves, the agent can take this decision for the user. The agent may then inform the user about this choice while noting which aspects of the situation led to this choice. This is a first step towards enabling the agent to explain its decisions to the user.

In this chapter we investigate the building blocks that would be needed to create such a personal assistant agent. First of all, we need a way to determine which meeting is considered to be more important to the user. To facilitate this process, we quantify the importance of each meeting by assigning it a numerical value to which we refer as the *priority* score of the meeting. Our assumption is that people implicitly follow this priority score when deciding about conflicting meetings by choosing the one with the highest priority. This will be evaluated via our research hypothesis:

RH - When choosing between two meetings, people select the one with higher priority in the majority of the cases.

The task of the agent now becomes to learn a model which predicts the numerical priority of meetings. We explore whether we can tackle this task by using machine learning techniques on a data set containing information on hypothetical meeting scenarios collected from multiple people. This leads to our first research question:

RQ1 - Can we use machine learning techniques to predict the priority of social situations based on situation cues and relationship features?

In our view for human-centered personal assistants, the ability of the agent to explain its decisions to the user is a fundamental requirement. This is because in such a system, it is important for the user to trust the suggestions of the agent. Lim et al. [103] suggest that in socio-technical applications, users trust the agent more when they understand why the agent has selected attending a specific meaning. Making the decisions of the agent explainable consists of three parts: the agent should be able to determine the internal processes that led to a certain suggestion, to generate an explanation based on them, and to present this explanation to the user [121]. Our focus is on the first part: we explore methods that allow the agent to determine which features of a social situation contribute to the prediction of priority. The other parts will be explored in future work.

Our predictive model is built using information from multiple people, however people can have differing preferences. To achieve more personalization, we extend the model by including personal values as input features . Values are considered to be a driving factor in human behaviour [53], so we explore their role in helping better predict the priority of social situations:

RQ2 - Does adding information about the personal values of users as input features to the predictive model increase that model's accuracy of prediction of the priority of social situations?

The rest of this chapter is structured as follows: In Section 4.2 we present our approach for tackling the research questions and hypothesis. In Section 4.3 we introduce background knowledge related to the concepts we use. Section 4.4 presents a crowdsourcing user study conducted to collect data for building and evaluating our models, which is done in Section 4.5. Section 4.6 concludes the chapter.

4.2. Proposed Approach

We propose an architecture that allows personal assistants agents to model the user's social situation, and use this information to predict the priority of this social situation, or, in other words, predict how important this specific situation is. A high-level depiction of the architecture is presented in Figure 4.1.

Overall, the framework works as follows: In the offline stage, a supervised learning algorithm takes as input multiple social situations from different users, described in terms of their social relationship features and situational features. The learning target is the priority of these situations. This forms our prediction model. During run time, the personal assistant agent is provided with the features of two different meetings which overlap. Using the priority prediction model, it determines the priority score of each meeting, it keeps on the schedule the meeting with the highest priority, and informs the user. At this point, the agent also determines the features that have the highest impact on this prediction, which will in future work lead to generating explanations.

In order to provide more personalized support, we add to the model information about the personal values of the user. The assumption is that people with similar value

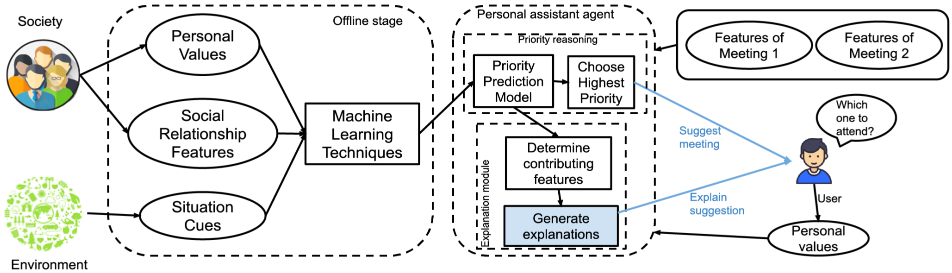

Figure 4.1: High level representation of the proposed architecture. Circles represent the modelled concepts, whereas boxes represent learning/reasoning steps. Items marked in blue are concepts that we do not explicitly tackle in this work. Icons used in Figure 4.1 were made by Freepik and retrieved from www.flaticon.com

4

preferences will also assign more similar priorities to specific situations. For instance, users who value achievement and success might give a higher priority to work meetings. Therefore, having the value information as an input in our model can potentially lead to better predictions.

A key concept that we use in this chapter is assigning a numerical score to the priority of social situations. Using this approach, as opposed to directly choosing between two conflicting meetings from a set of input features, has several advantages. First of all, using priority can facilitate the explanations given to the user by the personal assistant agent: the agent first tells the user which meeting has the highest priority, and secondly it explains why. Furthermore, having a numerical representation of priority comes with technical benefits, since the task of learning preference rankings from pairwise choices can be computationally intractable [28].

4.3. Concepts and Methods

In our vision, personal assistant agents should be able to provide human-centred support. This means the support actions should be transparent and intelligible. This guides our choices from two points of view: concepts (Section 4.3.1) and techniques (Section 4.3.2) used for modelling. This means we should use techniques that allow insight into their decision making process, combined with concepts that are understandable to the users. For this reason, we combine explainable machine learning techniques with concepts from social sciences. Using explainable machine learning techniques means that, when given a set of features which model a social situation, the model is able to output both a prediction as well as which features contributed to this prediction. Lim et al. [103] show that such a procedure improves the intelligibility of context-aware intelligent systems. The set of features that we use to model social situations is borrowed from social science literature. Since these are concepts that we use in everyday life, they should be understandable to the user. In this section, we present the rationale behind the concepts and techniques that we use.

4.3.1. Social Science Concepts

Our focus in this work is on modelling social situations - situations involving our user and people from their social circle. In Chapter 3 we propose modelling social situations involving two people as a combination of social relationship features, which represent how the two people are related to one another, and situation cues, which represent the circumstances in which the situation takes place. In this approach, the set of features that describes a social situation is based on social science literature that aims at modelling the relevant aspects of social relationships and situations. These concepts can be both concrete and objective (e.g., geographical distance between the user and the other person, for how long they know each-other), as well as subjective (e.g., quality of the relationship between the user and the other person).

Personal values represent key drivers of human decision making [143], [150]. Friedman and colleagues [53] define values as “what a person or group of people consider important in life". People hold various values (e.g. wealth, health, independence) with different degrees of importance. The most prominent models of human values were proposed by Rokeach [143] and Schwartz [150]. In our work, we use the model proposed by Schwartz since it offers validated measurement instruments with fewer items than Rokeach, which makes them more suited to online surveys. Furthermore, Schwartz builds on the work of Rokeach and other researchers, so there is overlap in their proposed value lists.

4.3.2. Machine Learning Methods

A predictive model for a personal assistant has to fulfil two main requirements. Firstly, it should be able to achieve satisfying accuracy for smaller data sets, since acquiring large amounts of data from human subjects can be challenging. Secondly, in order to provide human-centred support, the algorithms should provide insights on which features influence a specific prediction. Ensemble methods [34] are a family of machine learning techniques which fit these requirements. These techniques combine predictions from multiple learning algorithms in order to increase accuracy. The idea is to combine accurate and diverse weaker learners, in order to exploit the strength of each learner. Ensemble methods are shown to perform better than the individual learners they consist of [34]. Furthermore, they are shown to perform well and generalize better for smaller data sets [133]. When the base learners are decision trees, it is also possible to have insights into the features that led to a certain decision, as we will show further on. Another advantage is their accuracy when dealing with structured data. A recent survey [147] shows that ensemble methods have won different machine learning competitions, thus demonstrating high predictive power.

Some of the more successful methods are random forests [20] and gradient boosting machines [54]. Random forests are a specific example of bagging methods [19]. In bagging, each learner is built independently over a random sub-sample of the data, and the decision is made by aggregating the outputs. The sub-sampling procedure reduces the variance of the method, which is usually a problem for decision trees. In addition, in random forests, for each split of the tree only a sub-set of the features is considered, this way we avoids the possibility of all the trees selecting the same features and ignoring others. Gradient boosting machines are an example of boosting methods [54]. In boosting, weaker learners are trained sequentially (and not in parallel like in bagging), and each new learner tries to correct its predecessor. Gradient boosting achieves this by fitting the new predictor to the residual errors of the previous ones.

Understanding why a model makes certain predictions is a general goal in machine learning, especially when it comes to providing human-centred support. Lundberg and Lee [106] propose a unified framework for interpreting predictions, called Shapley Additive Explanations (SHAP). The benefits of using this framework are that it provides both global interpretability for the model (i.e. which are generally the most important predictors), as well as local interpretability (i.e. which are the most influential predictors for each individual observation). We use this framework in order to gain insight into the predictions of our model.

4.4. Crowd-sourcing User Study

In this user study we gather data for constructing and evaluating our models. The study was approved by the ethics committee of Delft University of Technology.

4.4.1. Choice of Concepts

Features of Social Situations

In order to set up the user study, we need to define a set of features that will be used to model social situations. Our starting point is the feature set proposed in Chapter 3, where social situations are described through a set of relationship features and a set of situation cues. Since their feature set is based on a limited number of social science models, we start by conducting a more extensive literature review. Then, we conduct an exploratory pre-study in order to investigate what aspects of a social situation people take into account when determining how important that situation is. Thus, our feature set is evaluated both from a theoretical and practical perspective.

In our literature review, we found five comprehensive models which aim at describing aspects of dyadic social relationships, three of which were not taken into consideration in Chapter 3. The results are presented in Table 4.1.

Next, in the exploratory pre-study, we collected answers from 33 participants through Amazon Mechanical Turk[1]. Our goal was to explore which features do participants find important when thinking about the priority of social situations. First, participants were asked to describe five social situations in which they participated in the past week. They were instructed to provide at least the time, location, activity and role of the other person, in order to ensure they were thinking about concrete situations. Furthermore, these suggested activities provide the basis for the formation of the hypothetical scenarios in our main user study (Section 4.4.2). Then, they were asked to consider which aspects of the situation play a role in determining the priority they would assign to a social situation, and the relative importance of these aspects towards determining priority. This question was asked separately for the relationship features and the situation cues. In both questions, participants were free to add answer options, and for each feature they ranked its importance (i.e., how much is it taken into account) in determining the priority of the situation on a 5-point Likert scale ranging from ‘Not at all important’ to ‘Ex-

[1] https://www.mturk.com/

tremely important'. The set of relationship features that had an importance of 3 or higher and were mentioned by at least 20% of the participants, are marked with a + in the last column of Table 4.1.

The relationship features that we use to model social situations are marked in bold in Table 4.1. We select the aspects which appear in at least two columns of the table. To that set, we add two more features, namely the age difference between the user and the other person, and whether the two have the same or different genders. Despite not being directly relationship features, age and gender appear as relevant aspects of social relationships in most research from social sciences [23], [131], so we believe this warrants their addition to our model. These features are not included in Table 4.1, since it exclusively contains relationship features.

Table 4.1: Different aspects of social relationships present in the literature as well as in the exploratory pre-study. The items written in bold text form our set of social relationship features. Items marked with an asterisk are the features proposed in Chapter 3.

Relationship Feature	[131]	[23]	[9]	[69]	[122]	Pre study
Role*	+	+	+	-	+	+
Contact Frequency*	+	+	+	-	+	+
Geo-distance*	+	-	+	-	-	+
Years known*	+	+	+	-	+	+
Hierarchy*	-	-	-	+	+	-
Relationship quality*	-	-	+	-	-	+
Depth of acquaintance*	+	+	-	+	-	+
Formality level*	-	-	-	+	+	-
Trust level*	-	-	-	-	-	-
Shared interests	+	+	-	-	+	+
Communication aspects	-	-	-	-	+	-
Reciprocity	-	-	-	+	-	-
Complexity	-	-	-	+	-	-

When it comes to situation cues, we use the ones proposed in Chapter 3, since the literature review and exploratory pre-study did not reveal new elements that warrant addition. Thus, our set of situation cues consists of: setting (work, family, sports, casual), event frequency (regular, occasional), initiator (user, other person, neither) and help dynamic (giving, receiving, neither).

PERSONAL VALUES

For a list of personal values to elicit from participants, we turn to the European Social Survey [151]. It consists of a list of 18 statements (two for each universal value group - Self-direction, Stimulation, Hedonism, Achievement, Power, Security, Conformity, Benevolence and Universalism) that describe features/qualities of a person (e.g., “Thinking up new ideas and being creative is important to him/her/them. He/She/They like(s) to do things in his/her/their own original way."), where each statement represents a personal value (e.g., creativity). The subjects were asked to assess how similar they believe this person is to them, on a scale from 1 (Not like me at all) to 6 (Very much like me). The original survey consists of 21 values, however, we removed the statements of the value group “Tradition" since its values (devotion, religion) do not fit with the type of scenarios that participants were presented with. Furthermore, in the category “Security" we replaced the statement for the value National Security with the statement for the value Health for the same reason. The statement for the value Health was taken from an extended version of this survey which consists of 40 items [151].

4.4.2. METHOD

PARTICIPANTS

4

We recruited 302 subjects on the online crowd-sourcing platform Prolific Academic. Using a crowd-sourcing platform allowed us to efficiently obtain a large sample size in a short amount of time. Respondents received monetary compensation for the time they spent, as per the platform policies. After eliminating the ones who did not pass at least two of our three attention checks, our data consists of answers from 278 subjects. 149 of them are female, 127 are male, and two participants selected the option “other" when asked about their gender. The average age of the subjects is 36.2 years old (SD=12.3).

PROCEDURE

Subjects answered an online survey[2]. After being briefed about the purpose of the study, they were presented with its four parts. In the *first part*, subjects were asked about their relationship with five people from their social circle. The questions were the relationship features that are marked in bold in Table 4.1. Ideally, we wanted the subjects to select people with whom they have different types of relationships. In Chapter 3 we suggested that when left without guidance, subjects tend to select people closer to them. This, in turn, leads to less variety and a more imbalanced data set. To avoid this, we predetermine some of the features as follows: the first person the subject selects had to be a family member. The second person had to be one of their (current or past) direct supervisors or managers. The third person had to be someone with whom they have a negative or very negative relationship. The fourth person had to be one of their friends. The last person had to be someone that the subject does not know very well. Subjects were instructed to simply provide us with the initials of these people. This way, on one hand anonymity is preserved, and on the other hand, we could refer back to these people in the next parts of the experiment.

In the *second part*, subjects were presented with eight hypothetical social situations, which were meeting scenarios involving one of the people from the first part (selected randomly). We used hypothetical situations, since this gives us control over the types of situations subjects are presented with, ensuring a wide variety. To make the situations seem realistic, we presented subjects with activities that are common for people in their daily lives. Meeting situations were formed by combining situation cues: setting, activity within setting, event frequency, initiator, and help dynamic, as described in Section 4.4.1 (E.g. “You have a weekly work meeting with your team leader where you expect

[2] The survey questions and the data can be found in the supplementary materials in `https://doi.org/10.4121/13176923`.

to get feedback on a project that you are working on.”). Activities are not part of our situation cues, however, we included them in the description of the scenarios in order to make them more concrete. These activities were collected in the exploratory pre-study described in Section 4.4.1. The activities were grouped into settings, and for each setting, we selected the ones that were suggested more often: four for the casual setting, three for the work setting, three for the family setting, and two from the sports setting, for a total of twelve activities. We selected more activities for the casual setting and less for sports, to reflect the proportions of activities mentioned by the participants of the exploratory user study. Each subject was presented with eight of these twelve activities. Subjects were asked what priority they would assign to each situation on a 7-point Likert scale (ranging from Very Low to Very High). Furthermore, they were asked how likely they are to encounter a similar situation in their daily life on a 5-point Likert scale (ranging from Very Unlikely to Very Likely). This information is used to assess whether the hypothetical scenarios seem realistic to the subjects.

In the *third part*, subjects were presented with five pairs of situations (from the second part), and for each pair, they were asked the following question: “Suppose that in a certain week you are very busy due to some other unexpected commitment, so you can attend only some meetings and cancel some others. Which of these two meetings would you attend?". Lastly, in the *fourth part* subjects answered the survey about personal values described in Section 4.4.1.

4.4.3. Description of Data

In order to be able to build a model that generalizes better, it is important to have a wide variety of data. Overall, we notice that this is the case for most of the social features. The roles were represented as follows: friends - 29.5%, family members - 26,31%, supervisors/managers - 21.3%, co-workers - 8.71%, neighbours - 5.53%, members of the same group - 3.02%. Features such as geographical distance (64.8% living less than 1 hour away), depth of acquaintance (mean=3.28, SD=1.33), frequency of contact (mean=2.91, SD=1.4) and formality level (mean=2.27, SD=1.45) follow a similar distribution to the ones reported in Chapter 3, so we do not report them fully for space purposes. Relationship quality was on average slightly positive (mean=0.55, SD=1.26, on a scale where -2=very negative, -1=negative, 0=neutral, 1=positive, 2=very positive). Fixing its value for one of the selected people led to more balanced answers for relationship quality as compared to the ones reported in Chapter 3.

When it comes to the priority of the scenarios, subjects assigned relatively high priorities. The average priority was 5.12 (on a 7-point Likert scale), with a standard deviation of 1.96. Participants found the scenarios on average to be relatively realistic (mean=3.02, SD=1.5, on a 5-points Likert scale), with 47.9% of the scenarios being ‘Likely’ or ‘Very Likely’.

In the third part of the user study, we asked subjects to specify which meeting they would attend if they had to select between two meetings. We use this data to test whether subjects mostly select meetings which have a higher priority. In 25% of the cases, subjects were presented with two meetings which have the same priority, so we cannot use this fraction of the data to test our hypothesis. This is an unintended result of the experimental setup, and in future experiments this can be controlled beforehand. For the data in which it is possible to make a distinction, subjects select the meeting with a higher priority in 58% of the cases, and the one with lower priority in 42% of the cases. This result marginally supports our research hypothesis, however, 42% remains a large figure. One potential reason can be the noise in the data caused by the fact that we present subjects with hypothetical scenarios, since some of these scenarios are situations that subjects do not normally encounter in their lives. To test this assumption, we remove the meetings which subjects consider to be 'somewhat unlikely' or 'very unlikely' in part 2 of the experiment. In the remaining data, in 68% of the cases subjects select the meeting with the higher priority. This is significantly higher than 58% (Two-Proportions Z-Test, $p < 0.05$), which suggests unlikely meetings can be a source of noise. Further reasons why some subjects select the meeting with a lower priority will be explored in future work.

When asked about personal values, subjects reported on average the following scores (on a 6-points scale): Benevolence - 4.81 (SD=0.93), Self-direction - 4.75 (SD=0.93), Universalism - 4.7 (SD=1), Security - 4.49 (SD=1), Hedonism - 4.03 (SD=1.09), Conformity - 3.96 (SD=1.22), Stimulation - 3.87 (SD=1.26), Achievement - 3.78 (SD=1.28), and Power - 3.22 (SD=1.36).

4.5. PREDICTING PRIORITY OF SOCIAL SITUATIONS

In the following subsections, we use the data from the crowd sourcing user study to explore our research questions.

4.5.1. PREDICTIVE MODELS AND RESULTS

In this section, we present the models that we use to predict the priority of social situations, and compare their performance. Models take as input the full list of social relationship features and situation cues. Subjects could assign priorities on a scale from 1 to 7, so we model this task as a regression task, since there would be too many classes to model it as a classification task for the amount of data that we have. This means, given a set of features, the model predicts a continuous score between 1 and 7. We believe this should not pose an issue although subjects were presented with discrete answer choices, since these choices were ordinal, and the concept of priority is in itself continuous.

As mentioned in Section 4.3, we use a random forest model as well as a gradient boosting machine model. Specifically, we use the RandomForestRegressor and XGBRegressor implementations from the Scikit-learn package in Python[3]. We split the data and randomly assign 80% to the training set and 20% to the test set. We perform parameter tuning by using cross validation on the training set. We report the performance of these models on the test set. For comparison we include a decision tree model, since this approach was previously used to predict the priorities of social situations in Chapter 3. Furthermore, we include three baseline predictors based on heuristics, namely: an algorithm which always predicts the mean priority score, an algorithm which predicts a random score between 1 and 7, and an algorithm which always predicts the most chosen class (in this case, a priority of 7). Including such baseline predictors is common practice for new machine learning tasks with no predetermined benchmarks (e.g. [66]).

[3]The code can be accessed under: https://github.com/ilir-kola/priority-social-situations.git

We start by reporting the Mean Absolute Errors, as well as the Mean Squared Errors for predictions on the test set. Results are reported in Table 4.2.

Table 4.2: Model errors in predicting the priorities of situations. Differences between predictions are statistically significant ($p < 0.05$). In bold, the best performing model.

Model	Mean Absolute Error	Mean Squared Error
Random Prediction	2.53 (SD=1.76)	9.17
Predict Most Chosen Class	1.84 (SD=1.93)	7.1
Predict Mean	1.56 (SD=1.15)	3.72
Decision Tree	1.81 (SD=1.79)	6.21
Random Forest	**1.35 (SD=1.02)**	**3.25**
XGBoost	1.43 (SD=1.12)	3.34

As we can see from the results, the best performing model is the Random Forest model, followed by XGBoost. They outperform the Decision Tree model, as well as the baseline heuristic predictors that we used as a comparison. In practical terms, it means our best model on average makes a prediction error of 1.35, on our 7 point scale. However, this number is just an average, so it gives limited insight into individual predictions. For this reason, we look more in detail into what does this error mean for the three best performing models from Table 4.2.

In general, our data set is to some extent unbalanced, since there are more situations which receive a high priority (i.e. 5, 6 or 7) compared to the ones receiving a low priority (i.e. 1, 2 or 3). In our specific domain - a personal assistant that manages the user's agenda - it is often more important to be able to distinguish a situation with a low priority from one with a high priority (or vice versa), rather than to be able to differentiate between two meetings with different degrees of high (or low) priority. This is a well-known controversy (e.g. [157]) arising from interpreting Likert scales as numeric intervals: a prediction error of 2 which confuses 'Slightly high' with 'Very high' does not have the same nuance as a prediction error of 2 which confuses 'Slightly high' with 'Slightly low', because of the change of category (from high to low) involved in the latter example. By dichotomizing our data into situations with high priority (i.e. with a priority higher than 4) and low priority (i.e. with a priority lower than 4), we can evaluate how often do the predictors assign a high priority to a situation with a low priority, as well as the other way around (similar to Type 1 and Type 2 errors). The algorithm which always predicts the mean (i.e., 5.12) always predicts a high priority, so it is always right for situations with a high priority, and always wrong for situations with a low priority. The Random Forest model and XGBoost perform equally well for high priority situations: none of them is classified to have a low priority by Random Forest, and only 2.17% of them by XGBoost. When it comes to situations with low priority, these models clearly outperform the heuristic predictor: Random Forest wrongly classifies only 30% of situations to have high priority, whereas for XGBoost the value is 29.5%.

Our results suggest that Random Forest and XGBoost outperform heuristic predictors both in absolute errors as well as when considered in the context of our application domain. Random Forest has a slight edge on XGBoost, however, the difference is not high enough so as to declare a clear winner.

4.5.2. DETERMINING IMPORTANT FEATURES FOR PREDICTIONS

A key advantage of the machine learning models is the fact that it is possible to get insight into their decision process. This allows for the possibility to explain to the users which features led to a certain prediction, and adapt the model if needed. We use the TreeExplainer method of the SHAP package, which is based on the work of Lundberg and Lee [106].

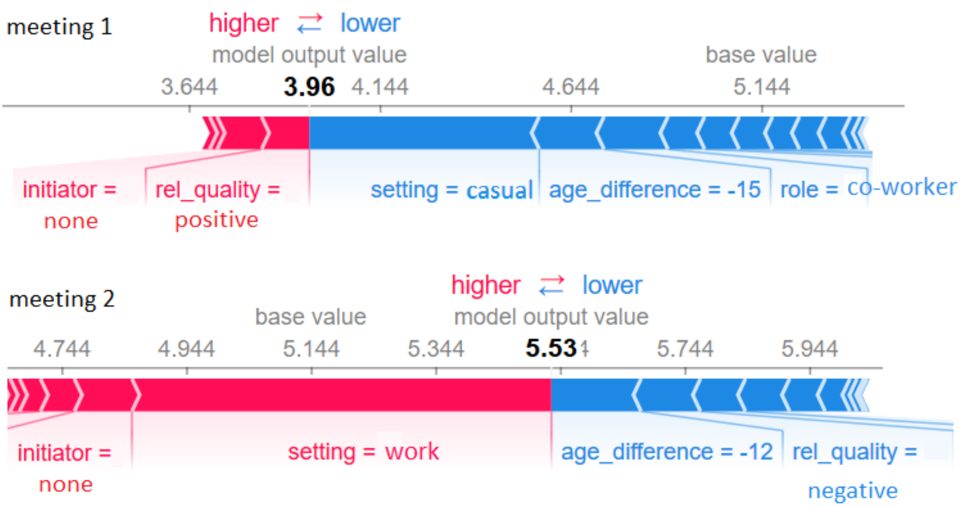

Figure 4.2: Explaining the features that led to specific predictions. Larger bars have more impact on the decision. Features marked in red contributed to making the priority prediction higher, whereas the ones in blue lower. The text under the bar indicates the values of each feature for the specific situations.

From a global perspective, the most informative features are setting, relationship quality, age difference, and role. This means these are the features that mostly contribute to the predictions. However, when running the predictive model without the least important features (i.e., hierarchy, geographical distance, other person's gender), we notice a drop in accuracy. This suggests that all the features are to some extent important in predicting specific situations.

To illustrate the interpretations of individual predictions we use two specific social situations which our model had to predict (Figure 4.2). In both situations, setting is the feature with the highest influence. We notice in this example that the work setting causes the meeting to have a higher priority, whereas the casual setting contributes to a lower one. As expected, a positive relationship quality makes the priority of the meeting higher, as opposed to the negative relationship quality. In both cases, meeting a younger person contributes to a lower priority.

This method allows for insight into the decision process of the agent, and can form a basis towards explaining the suggestion to the user. Miller [115] proposes that explanations in AI should be contrastive: people want to know why the agent suggested a certain action rather than another one. This is inherently part of our method, since the agent can explain to the user why one meeting was selected instead of another. Furthermore, people prefer an explanation that consists of a few causes rather than many. Using the SHAP package allows this, since it identifies the features with the highest impact.

4.5.3. ROLE OF PERSONAL VALUES IN PREDICTING PRIORITIES

We start our work by building a predictive model using data from multiple people, however, we want to explore whether it is possible to have some degree of personalization for the user. To achieve this, we explore whether adding information on the personal values of the users helps to increase the accuracy of the model. The underlying assumption is that users with similar value preferences will assign similar priorities to situations. This is based on the definition of values, which are considered to be drivers of behaviour. First, we train our Random Forest model with the original set of features, as well as 9 new features representing the score that the user assigned to each of value groups (Section 4.4.1), collected in the last part of the user study. The mean absolute error, in this case, is 1.38. This means that the quality of the predictions slightly deteriorates when adding information about values. One reason for this might be that adding 9 new features to the existing ones causes the model to have too many features, which can deteriorate performance. Another possible reason can be related to the salience of personal values in different situations. Schwartz [150] argues that in order for values to influence action not only should they be important to the actor, but they should also be relevant in that specific context. Kayal et al. also propose the use of domain values in order to reason about social commitments [83]. We check for this insight in our data. Some situations explicitly mention that the user is expected to help someone. Therefore, the value 'helpfulness' is salient in these situations. We notice that subjects who value helpfulness more, on average assign a significantly higher priority to situations where they have to help someone, as compared to subjects who value helpfulness less (6 vs. 5.09, $p < 0.01$ when performing the Mann-Whitney test). For meetings that do not involve giving help, the differences in the priorities assigned by these subjects are not significant. This suggests that certain values which are salient to the domain can potentially help predictions.

4.6. CONCLUSION

4.6.1. CONTRIBUTIONS

In this chapter, we propose an approach which enables personal assistant agents to predict the priority of a user's social situation. The approach relies on concepts from social sciences which are used to model social situations, as well as machine learning techniques which are used to learn the priority scores from data from multiple people. First, we review literature from social sciences and propose a set of features which we use to model the social situations of a user. Then, we conduct a crowd-sourcing user study in order to gather the data needed to build our predictive models and evaluate our approach. The subjects' answers suggest that having a numerical representation of priority can in principle be used to help deciding which meeting to attend in cases of overlapping meetings. The results marginally supports our hypothesis (RH): 58% of the subjects select the meeting with a higher priority. This can form the basis for allowing a personal assistant agent to use its priority predictions to choose between the meetings.

Next, we show that ensemble models such as Random Forests outperform baseline models in predicting the priorities of social situations, especially when it comes to differentiating between situations with high and low priorities (RQ1). Furthermore, we present a procedure which enables the personal assistant agent to determine the features that contributed to the predictions, which in future work will be presented as explanations to the user. We envision that this, together with the fact that features are taken from social science literature and are therefore more understandable for people, can help achieving the vision for more transparent and intelligible personal assistant agents. Lastly, we test whether adding information about the personal values of the user can help us lower the prediction error (RQ2). Results show that in our setting this is not the case, since the mean absolute error of the model suffers a slight increase. However, insights from the data suggest that using personal values which are salient in the specific situation has the potential to be a more successful approach.

4.6.2. LIMITATIONS AND FUTURE WORK

First of all, our experimental setup presented subjects with hypothetical scenarios. This was done to ensure variety in the data, however, this comes at the cost of the data being noisier, since some of the scenarios might be unlikely to actually occur, so the subjects might not answer consistently. It would be useful to conduct a user study in which subjects report all their social situations from a fixed period of time, in order to evaluate our models with more realistic data. Furthermore, asking subjects which meeting they would attend when two meetings overlap (part 3 of the user study) presented them with a binary choice, which does not inform us how certain they were about their selection. An alternative would be to provide participants with a slider, where they can state how inclined they would be to attend one of the meetings [83]. In future work, we aim to enable personal assistant agents to provide full explanations regarding their decisions to the user. This is based on our assumption that presenting the user with the social features that contributed to a prediction makes the work of the personal assistant agent more transparent. This has to be tested in practice. Next, we will investigate the possibility of a feedback loop between the user and the agent based on the explanations, in order to further personalize support. This would be important especially for cases where the subjects disagree with the agent's decisions. Lastly, we will explore whether our models can be used to predict other aspects of social situations other than priority.

5

Social Situation Comprehension through Psychological Characteristics of Situations

Support agents that help users in their daily lives need to take into account not only the user's characteristics, but also the social situation of the user. Existing work on including social context uses some type of situation cue as an input to information processing techniques in order to assess the expected behavior of the user. However, research shows that it is important to also determine the meaning *of a situation, a step which we refer to as social situation comprehension. We propose using psychological characteristics of situations, which have been proposed in social science for ascribing meaning to situations, as the basis for social situation comprehension. Using data from user studies, we evaluate this proposal from two perspectives. First, from a technical perspective, we show that psychological characteristics of situations can be used as input to predict the priority of social situations, and that psychological characteristics of situations can be predicted from the features of a social situation. Second, we investigate the role of the comprehension step in human-machine meaning making. We show that psychological characteristics can be successfully used as a basis for explanations given to users about the decisions of an agenda management personal assistant agent.*

This chapter is currently under review [93].

5.1. INTRODUCTION

Artificial agents that support people in their daily lives – such as personal assistants, health coaches, or habit formation support agents – are becoming part of everyday lives (e.g. [84], [132]). Existing work on personal agents usually focuses on modelling personal characteristics of the user, such as their goals, emotional state, or personal values (e.g. [31], [51], [97]). However, research in social science shows that human behavior is not only shaped by a person's state and characteristics, but also by the situation they are in [101]. This suggests that in order to provide better aligned support, personal agents should take the user's situation into account in determining which support to provide.

In this chapter we take a step towards addressing this challenge, with a specific focus on the *social* dimension of situations. This is important because our daily situations often have a social nature: we spend time at work with colleagues, and free time with family and friends. Support agents thus need to account for the social dimension of situations, and how that affects the behavior of users. The need for enabling support agents to understand the social situation of the user has been acknowledged as an important open question in agent research [158], [166]. More broadly, the ability to assess and act in social situations has been proposed as the next challenge that intelligent machines should tackle [65].

5

5.1.1. MOTIVATION

Existing approaches (e.g., [2], [38], [96]) tackle this challenge by using some type of situation cues as input (e.g., actors, relationship characteristics, etc.), and using information processing techniques such as machine learning or rule-based approaches to assess expected behavior. By going directly from social situation features to predicted or desired user behavior, the step of understanding the *meaning* of the social situation from the point of view of the user is not performed explicitly. However, research in social psychology (e.g., [44]) suggests that people determine how to behave in a situation by ascribing meaning to this situation, and using this interpretation to decide how to act.

Inspired by this insight, in Chapter 2 we propose that support agents should perform this step explicitly. They refer to this process as *social situation comprehension*. Following research on situation awareness [47], they propose a three-level architecture where social situation comprehension is the middle level (Level 2) in between social situation perception (Level 1) and social situation projection (Level 3), as depicted in Figure 5.1. The idea is that Level 2 information is derived from Level 1 information, i.e., social situation features, and Level 3 information about expected user behavior is in turn derived from Level 2 information.

A central question in realizing such a three-level architecture is in what 'terms' the meaning of a situation should be described. In this chapter we investigate whether *psychological characteristics of situations*, a concept used in social psychology (e.g., [126], [135], [179]), can be used for this purpose of achieving social situation comprehension in support agents. The idea behind psychological characteristics of situations is that people view situations as real entities, and ascribe to them traits or characteristics in the same way they ascribe characteristics to other people. For instance, the situation 'having a progress meeting with your supervisor' can have a high level of duty and intellect and a low level of deception and adversity. An important advantage of using psychological characteristics of situations is that they are general enough to model arbitrary daily life situations [135].

Our goal is to explore whether incorporating information about the psychological characteristics of the user's situation would be beneficial for support agents. Support agents should make accurate suggestions that are trusted by the user. We investigate the use of psychological characteristics in support agents from these two perspectives. First, we study whether they can be used for predicting user behavior (Level 3 information), which is a basis for accurate suggestions. Second, we investigate whether they can provide meaningful reasons for explaining the suggestions of the support agent to the user, since research [115] suggests that explainability of Artificial Intelligence (AI) systems is important for enhancing their understanding and in turn trustworthiness.

5.1.2. USE CASE

In this chapter we take the example of a socially aware agenda management agent. Our goal is not to build a socially aware agenda management agent in itself, but this use case has characteristics that make it ideal for exploring the effects of incorporating psychological characteristics of situations. First of all, making accurate predictions on which to base its suggestions and giving insightful explanations is crucial for this agent, which is in line with aspects we aim to explore. Secondly, through this case we can study future situations for which the information is available beforehand. This way, we can focus on how the information can be processed to interpret the social situation and its effect on user behavior rather than having to deal with run-time situation perception, since that is beyond the purpose of our current work. Furthermore, such an agent facilitates conducting online user studies since it allows us to frame social situations as meetings, an easy concept to explain to participants. Lastly, the types of possible meetings can be arbitrary rather than about a specific domain, thus allowing us to explore a wide variety of social situations.

Providing support to the user regarding which meeting to attend can be seen as choice support. According to Jameson et al. [76], in choice support the goal is to help the chooser (i.e., the user) make the choice in such a way that, from some relevant perspective, the chooser will be satisfied with the choice. Jameson et al. [76] present different choice patterns that people tend to follow and how technologies can support people in these choices: Access information and experience, Represent the choice situation, Combine and compute, Advise about processing, Design the domain and Evaluate on behalf of the chooser. The agenda management agent used throughout the chapter gives suggestions to the users on which meetings to attend, thus following the 'Evaluate on behalf of the chooser' choice support pattern.

5.1.3. RESEARCH QUESTIONS AND HYPOTHESIS

An important aspect of agenda management is dealing with scheduling conflicts where not all desired meetings can be attended. We develop predictive models that would allow such an agent to determine the priority level of each meeting, taking into account its social aspects. This is done via determining the situation profile of each meeting consisting of the psychological characteristics of the situation based on the DIAMONDS model [135]. For example, dinner with a friend might be characterized by a low level of duty, but high level of positivity and sociality, while a meeting with a difficult colleague at work might be characterized by a high level of duty, high use of intellect and high level of adversity. This information is used to determine the priority level of each meeting, which is expected to correspond with the user behavior of choosing a high priority meeting in case of scheduling conflicts. The agent would make a suggestion to the user about which meeting to attend.

Based on this description, we formulate the following research hypothesis:

> **RH** - Using psychological characteristics of a social situation as input in a machine learning model leads to a more accurate prediction of the priority of the social situation than using social situation features as input.

Collecting information about the psychological characteristics of each situation would be an intrusive task, therefore in the next research questions we explore whether we can automatically predict the psychological characteristics of a situation, and how useful would these predictions be:

5

- **RQ1** - To what extent can we use machine learning techniques to predict the psychological characteristics of a social situation using social situation features as input?

- **RQ2** - To what extent can we use the predicted psychological characteristics from RQ1 as input in a machine learning model to predict the priority of a social situation?

Since we use explainable techniques for creating the predictive models, this also allows to determine which features were the most salient in determining the priority. These can be presented to the user as explanations. Following the previous example, if the two meetings are overlapping the predictive model might determine that the second meeting is more important and that the most salient feature is duty. In that case, the agent would tell the user '*You should attend the second meeting since it involves a higher level of duty, and meetings with higher level of duty are usually prioritized*'. Through the following research questions we explore the perceived quality of such explanations:

- **RQ3** - To what extent can social situation features and psychological characteristics of situations be used as a basis for explanations that are complete, satisfying, in line with how users reason, and persuasive?

- **RQ4** - When do people prefer psychological characteristics of situations in explanations compared to social situation features?

Our work has an exploratory nature, since the topic of incorporating psychological characteristics of situations in support agents is novel. For this reason, we do not always have a preconceived idea of the relation between variables to form hypotheses. Posing research questions allows us to explore and provide initial insights on the topic without being bound to specific expected outcomes. We assess these questions through two studies, one which addresses the predictive powers of psychological characteristics by creating machine learning models, and one which performs a user study to investigate the use of different kinds of explanations. The rest of the article is organized as follows: Section 5.2 gives an overview of background concepts that we use throughout the chapter. Section 5.3 introduces the first study, presents and discusses its results, and addresses **RH, RQ1** and **RQ2**. Section 5.4 introduces the second study, analyzes and discusses its results, and addresses **RQ3** and **RQ4**. Section 5.5 concludes the article.

5.2. BACKGROUND

This section positions this chapter in relation to existing work and offers an overview of background concepts that are used throughout the chapter. In particular, we present the three-level social situation awareness architecture proposed in Chapter 2 which forms the starting point for our work.

5.2.1. RELATED WORK

The concept of sociality is broad, and so are its applications to artificial agents. The main directions involve agents being social with other artificial agents, and agents understanding human sociality. The agent technology research community has explored sociality from the point of view of artificial agents interacting with each other in multi-agent systems governed by structures such as norms, institutions and organizations (e.g., [37], [52], [107]). The other research direction explores the sociality of agents in relation to humans. This is researched from the perspective of agents interacting socially with people (e.g., [32], [45], [165]), and agents modelling human sociality. An example of the latter is research on social signal processing, which focuses on using social cues such as body language to assess behavior [167]. Other approaches more closely related to ours employ some type of social situation information as input, and process that information to assess expected user or agent behavior. In our work we take inspiration from the way in which they conceptualize social situations. The key difference is that we explicitly reason about the meaning of the social situation for the user.

Dignum and Dignum [38] propose using social practices [137]. Social practices are seen as ways to act in context: once a practice is identified, people use that to determine what action to follow. For instance, the social practice 'going to work' can incorporate the usual means of transport that can be used, timing constraints, weather and traffic conditions, etc. A social practice is identified using information from physical context, social context, activities, etc. Social context includes information about places and roles. Each social practice contains a concrete plan which makes the connection between the social context input and the behavior that needs to be manifested in that situation.

Ajmeri et al. [2] also highlight the importance of modelling social context in personal agents. Social context includes information such as the place of the interaction or the social relationships between the people in the interaction (i.e., their role). In their approach, the agent includes the social information in the form of norms and sanctions that guide the agent's behavior. These norms and sanctions are formalized as rules in which the social context information serves as the antecedent and the behavior serves as the consequent: the agent exhibits a specific behavior only in presence of specific social context information.

Another approach on how to take into account the effects of social situations on user behavior is proposed in Chapter 4. They model social situations through a set of *social situation features* seen from the point of view of the user. For instance, in a situation where a manager and an employee are meeting, the support agent of the employee would model this situation through features such as *setting=work, role of other person=manager, hierarchy level=higher* and so on. Different from the previous approaches, in this work the relation between the social situation information and the expected behavior is learned rather than modelled explicitly. The authors show that it is possible to use these social situation features as input to a machine learning model to predict expected behavior such as the priority that people would assign to different social situations.

5.2.2. SOCIAL SITUATION AWARENESS IN SUPPORT AGENTS

This chapter builds on Chapter 2, where we propose a three-level architecture for social situation awareness in support agents. We define social situation awareness as: *"A support agent's ability to perceive the social elements of a situation, to comprehend their meaning, and to infer their effect on the behavior of the user"*. This definition instantiates Endsley's three-level model of situation awareness [47], yielding three corresponding levels of social situation awareness: social situation perception, social situation comprehension, and social situation projection. The resulting architecture is shown in Figure 5.1. The focus of this chapter is on the second level.

5

As can be seen from Figure 5.1, one of the key parts of situation comprehension is the ability to use Level 1 information for deriving a situation profile at Level 2. A situation profile is intended to express the meaning of the situation for the user. Level 1 information concerns features that describe salient aspects of the social situation. This information can come via sensory input or interaction with the user.

In Chapters 3 and 4 we propose a set of features based on research from social sciences. We divide features into situation cues, namely *setting, event frequency, initiator, help dynamic*, and social background features describing the social relation between the user and other people in the social situation, namely *role, hierarchy level, contact frequency, geographical distance, years known, relationship quality, depth of acquaintance, formality level and shared interests.* In the rest of this chapter we refer to these features as *social situation features* or *Level 1 information.*

The idea is that Level 1 information can be used to infer the meaning of the situation for the user, i.e,. Level 2 information. In this chapter we investigate the use of psychological characteristics of situations to model Level 2. As proposed in social science research, psychological characteristics of situations are used by people to ascribe meaning to a situation [135]. People use these psychological characteristics to predict what will happen in a situation, and coordinate their behavior accordingly. There are five main taxonomies which provide a set of psychological characteristics to describe situations [22], [57], [126], [135], [179], and in this work we use the psychological characteristics proposed in the DIAMONDS taxonomy [135]. This taxonomy has several advantages. Firstly, it is intended to cover arbitrary situations, and it offers a validated scale for measuring psychological characteristics. Furthermore, it is shown that the psychological characteristics of a situation correlate both with the features of that situation and

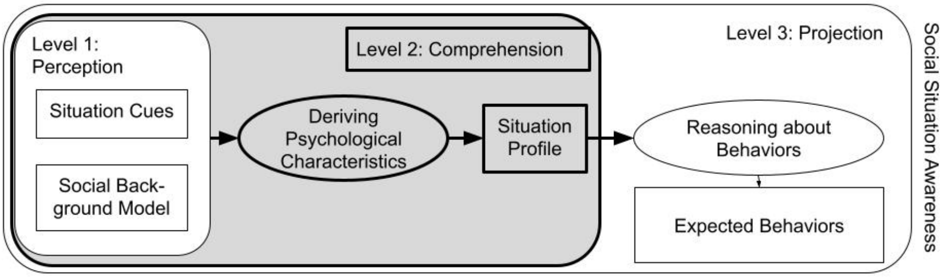

Figure 5.1: Simplified version of the three-level architecture for Social Situation Awareness proposed in Chapter 2 (emphasis on Level 2 added).

with the behavior people exhibit in that situation. The DIAMONDS taxonomy suggests that each situation can be described based on how characteristic each of the following concepts is:

- **Duty** - situations where a job has to be done, minor details are important, and rational thinking is called for;
- **Intellect** - situations that afford an opportunity to demonstrate intellectual capacity;
- **Adversity** - situations where you or someone else are (potentially) being criticized, blamed, or under threat;
- **Mating** - situations where potential romantic partners are present, and physical attractiveness is relevant;
- **pOsitivity** - playful and enjoyable situations, which are simple and clear-cut;
- **Negativity** - stressful, frustrating, and anxiety-inducing situations;
- **Deception** - situations where someone might be deceitful. These situations may cause feelings of hostility;
- **Sociality** - situations where social interaction is possible, and close personal relationships are present or have the potential to develop.

We call such a description a situation profile. In the rest of this chapter we also refer to the psychological characteristics of situations as *Level 2 information*.

The idea is then that a situation profile can be used by a support agent to determine expected behaviors for the user (*Level 3 information*), since research on the DIAMONDS model shows that there is a correlation between psychological characteristics of a situation and people's behavior in that situation. Information about expected behavior can in turn be used to determine how best to support the user.

5.2.3. EXPLAINABLE AI

Following the definition of Miller [115], when talking about explainable AI we refer to an agent revealing the underlying causes to its decision making processes. Early examples of such work can be found already more than forty years ago (e.g., [153]). In the last five years, this field of research has received increasingly more attention[1]. This is due to the increased availability of AI systems, as well as due to the emphasis on the importance of explainable AI coming from different governmental agencies [62], [68]. Different approaches have been proposed for explainable and interpretable AI (for an extensive survey, see [116]), and here we only provide a brief summary. Explanations can be global, i.e., explain the working of a system in general, and local, i.e., explain the reasons behind a specific decision or suggestion. Making the decisions of the agent explainable consists of three parts: the agent should be able to determine the internal processes that led to a certain suggestion, to generate an explanation based on them, and to present this explanation to the user [121]. Different techniques have been proposed to determine the internal processes of so-called black box algorithms (for a survey, see [67]). When it comes to the content of explanations, research shows that shorter explanations explaining why a certain decision (rather than another decision) is made are preferred [103], [115]. Furthermore, Ribera and Lapedriza [138] argue that explanations should be designed based on who the end user will be, and that explanations designed for lay users should be brief, use plain language, and should be evaluated via satisfaction questionnaires. We use these insights when designing the explanations for our user study.

5.3. STUDY 1 - PREDICTIVE ROLE OF PSYCHOLOGICAL CHARACTERISTICS

Through this study we evaluate our research hypothesis (**RH**), as well as **RQ1** and **RQ2**, as shown in Figure 5.2.

5.3.1. METHOD

In the first study we investigate to what extent psychological characteristics of situations can be used for predicting priority of meetings. Following the architecture in Figure 5.1, a situation profile (Level 2) should be derived from Level 1 information, and it should be able to predict Level 3 information. In order to create corresponding predictive models, we use data from a user study that collects information at Level 1 (social situation features), Level 2 (psychological characteristics) and Level 3 (priority) for a range of meeting scenarios.

The data that we use for building the predictive models was collected through the experiment described in Chapter 4[2]. The experiment was approved by the ethics committee of the university. Subjects were presented with meeting scenarios with people from their social circle (Level 1 information) and were asked to rate the psychological characteristics (Level 2 information) and priority of the meetings (Level 3 information).

[1] Google Scholar finds more than 22'000 publications in the time frame 2017-2022 for the search terms 'explainable AI', which is more than the number of publications for the time frame 1955-2016.

[2] The survey questions, the data and the source code can be accessed in the supplementary materials in https://doi.org/10.4121/16803889.

In Chapter 4 we use only part of the collected dataset which involves the social situation features (see Section 5.2.2) and the priority of hypothetical social situations. In this chapter we also make use of information about the psychological characteristics of each of the hypothetical social situations. First, to assess whether priority could in principle be predicted from psychological characteristics of situations, we take the 'true' Level 2 information as provided by our study participants, and create from this a predictive model for meeting priority (**RH**, top part of Figure 5.2). While this allows to assess the possibility to predict Level 3 from Level 2, our agent would not have the 'true' Level 2 information since it would be very cumbersome to ask users to provide this information for each meeting. This would not be the case for Level 1 information, since the social relationship features can be collected beforehand and tend to stay stable across situations. Thus, we want to investigate (see bottom part of Figure 5.2) whether we can predict Level 2 information from Level 1 (**RQ1**), and in turn, use these predicted psychological characteristics as input to predict Level 3 information (**RQ2**) using the predictive model that was built to assess our **RH**.

Data collection is a well-known obstacle when creating data-driven human decision predictive models. Using an experimental approach for collecting data is a good alternative when collecting data in the wild is not possible [145]. Furthermore, such an experimental approach can allow for more flexibility in the type of data that is collected. In the data set that we are using, the experimental setup presents participants with hypothetical meeting situations involving real people from their social circle. These hypothetical meetings are highly diverse in terms of their priority level and relationship features of the participant and the other person, including situations work meetings with supervisors, family occasions, casual meetings with friends etc. Explicitly capturing every aspect that is involved in how the user assigns a priority level to the meeting is not possible in practice for such a wide variety of meetings. Therefore, our goal is to explore whether modelling psychological characteristics of the situations can provide a good approximation that leads to accurate predictions of the priority levels.

5

Material

Social situation features used in the study were based on literature from social science (see Section 5.2.2 and Chapter 4). Specifically, the features used were: role of the other person, their hierarchy level, the quality of their relationship, the contact frequency, how long they have known each other, the geographical distance, the depth of acquaintance, the level of formality of the relationship, and the amount of shared interests.

Psychological characteristics of situations were taken from the DIAMONDS taxonomy (see Section 5.2.2), namely Duty, Intellect, Adversity, Mating, Positivity, Negativity, Deception and Sociality.

Scenarios used in this work represent social meeting settings that a user might encounter in their daily life. The scenarios had a hypothetical nature. Using hypothetical situations gives control over the types of situations subjects are presented with, ensuring a wide variety. To make these hypothetical situations more realistic, subjects were presented with activities that are common for people in their daily lives. Meeting situations were based on inputs from the users of a pre-study, and were formed as a combination of situation specific features (see Section 5.2.2): setting in which the meeting is taking

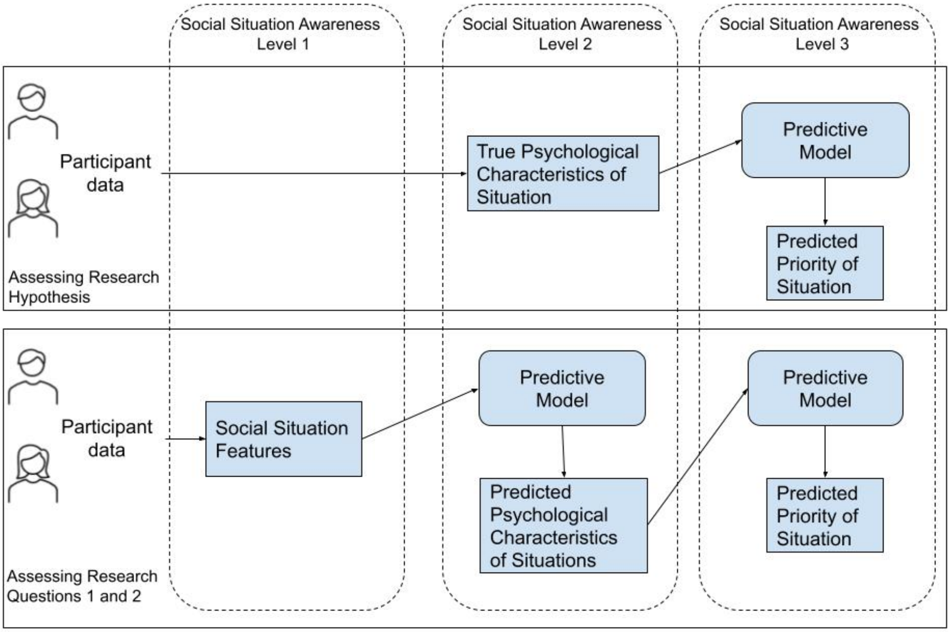

Figure 5.2: Conceptualization of Study 1, used to assess the research hypothesis (top part), and Research Questions 1 and 2 (bottom part)

place, frequency of meeting, initiator, and whether the user is expected to give or receive help (E.g. "You have a weekly meeting with AB[3] where you expect to get feedback on a project that you are working on."). In the situation descriptions, the setting was represented through typical activities that take places within that setting, to make the scenarios more concrete. For instance, the settings 'work' and 'casual' were represented by activities such as 'having a meeting with the supervisor' and 'going for dinner with a friend' respectively.

PARTICIPANTS

The study involved 278 subjects recruited through Prolific Academic[4], a crowd-sourcing platform where researchers can post studies and recruit participants who earn a monetary compensation for the time invested in conducting the study. 149 subjects were female, 127 were male, and 2 subjects selected the option 'other'. The mean age was 36.2, with a standard deviation of 12.3.

PROCEDURE

Subjects answered an online survey. First, participants were briefed about the purpose of the study. The goal of the study as conveyed to the participants was to collect information about the user's social relationships with different people from their social circle,

[3] For privacy reasons, users provided only the initials of people from their social circles.
[4] https://www.prolific.co/

as well as information about social situations involving the user and those people. Then they were presented with the two parts of the study.

In the first part, subjects were asked to select five people from their social circle, and then were asked questions about their relationship with these people using the set of relationship specific features (see Section 5.3.1). In the second part, subjects were presented with eight hypothetical social situations (see Section 5.3.1), which were meeting scenarios between them and one of the people that they mentioned in the first part of the study (selected randomly). Subjects were asked what priority they would assign to each situation on a 7-point Likert scale (ranging from Very Low to Very High).

Furthermore, subjects were asked about the psychological characteristics of each social situation using the dimensions proposed in the DIAMONDS taxonomy [135] (see Section 5.3.1). Subjects were presented with a description of each psychological characteristic, and they were asked “How characteristic are each of the following concepts for this situation?". Subjects answered on a 6-point Likert scale, ranging from Very Uncharacteristic to Very Characteristic.

In total, the dataset consists of information about 1390 social relationships between the subjects and people from their social circle, and about the priority level and psychological characteristics of 2224 hypothetical social situations involving the subjects and one of these people.

5.3.2. Results

The collected data is used to build predictive models which will be presented and evaluated in this section.

Using Psychological Characteristics of Situations to Predict the Priority of Social Situations

The task of predicting the priority of social situations was previously explored Chapter 4. There, we tested different learning algorithms that took as input the features of a social situation to predict the priority of that situation. If we refer to the social situation awareness architecture, that chapter takes as input Level 1 information and predicts Level 3 information. The best performing model was random forest, which led to a mean absolute error of 1.35, on a 7-points Likert scale.

For this reason, in this chapter we also employ a random forest model for predicting priority. The model takes as input the psychological characteristics of a social situation (Level 2), as obtained via the procedure described in the previous section, and predicts the priority of that social situation (as shown in Figure 5.2, top). Specifically, we use the RandomForestRegressor implementation from the Scikit-learn package in Python. We split the data and randomly assign 80% to the training set and 20% to the test set. We perform parameter tuning by using cross validation on the training set.

The results show that in our model, the mean absolute error is 0.98, which is a significant improvement (Wilcoxon Rank sum test, $p < 0.05$) over the 1.35 mean absolute error reported in Chapter 4. This suggests that psychological characteristics of situations are a better predictor of the priority of social situations than social situation features, thus supporting our hypothesis (**RH**).

PREDICTING THE PSYCHOLOGICAL CHARACTERISTICS OF SOCIAL SITUATIONS

The social situation awareness architecture introduced in Chapter 2, says that Level 2 information should be derived from Level 1 information. This is because having the agent ask the users about the psychological characteristics of each situation they encounter would be too invasive and time consuming. On the other hand, collecting Level 1 information can be done more efficiently, since the information about the social relationship can be collected in advance. For this reason, we investigate whether it is possible to predict the psychological characteristics of a social situation using as input social situation features (see Figure 5.2, bottom).

We evaluate the predictions of different regression algorithms: decision tree, XGBoost, Random Forest and Multi Layer Perceptor (MLP) using the scikit-learn library in Python. We train the models on 80% of the data, and evaluate them on the remaining 20%. We built 8 distinct models, where each model predicts one psychological characteristic, since this approach led to better accuracy than having one model that predicts all psychological characteristics at the same time. The model predicts a number from 1 to 6 (on a 6 point Likert scale, 1 being Very uncharacteristics, and 6 being Very characteristic), and the mean absolute errors are reported in Table 5.1. From the table (column 'Random Forest') we can see that, for instance, the model is on average 1.17 off when predicting the level of Intellect for a social situation. This means that for instance, if the real value is 5 (i.e. Moderately characteristic), the model is expected to predict a value between 3.83 (i.e. Slightly characteristic) and 6 (i.e. Very characteristic).

5

Table 5.1: Mean Absolute Errors of the models in predicting the psychological characteristics of situations. Psychological characteristics marked with * represent statistically different results between the best performing model and the model that predicts the mean (Wilcoxon Rank sum test, $p < 0.05$).

Psychological Characteristic	Decision Tree	XGBoost	Random Forest	MLP	Predict Mean
Duty*	1.66	1.36	1.34	1.38	1.55
Intellect*	1.48	1.21	1.17	1.23	1.3
Adversity	1.55	1.29	1.29	1.31	1.36
Mating*	0.92	0.87	0.85	0.93	1.03
Positivity*	1.44	1.18	1.14	1.17	1.26
Negativity*	1.51	1.25	1.25	1.39	1.37
Deception	1.14	1.04	1.04	1.09	1.09
Sociality*	1.42	1.06	1.02	1.03	1.13

In order to assess how good these predictions are, we compare our models with a heuristic model that always predicts the mean of the psychological characteristics. The results are reported in Table 5.1 (column 'Predict Mean'). We see that the random forest model significantly outperforms the heuristic predictor for all psychological characteristics apart from Adversity and Deception and always performs at least as well as the other predictive models. We use a heuristic model for comparison since this is the first benchmark result in predicting the psychological characteristics of a situation. Therefore we do not have an existing baseline to compare it with. Including heuristic baseline predictors is common practice for new machine learning tasks with no predetermined benchmarks (e.g. [66]). In Chapter 4 we also use heuristic predictors as a baseline for priority prediction, and the most accurate heuristic in that work is an algorithm that always predicts the mean priority.

In the next section we evaluate whether these predictions are sufficiently accurate to be used as an intermediate step for predicting priority of social situations. This allows the evaluation of the usefulness this predictive model as part of the bigger social situation awareness architecture.

Predicting Priority through Predicted Psychological Characteristics

To assess the usefulness of these predicted values for predicting the priority of social situations, we predict priority by using as input not the 'true' psychological characteristics of the situation as reported by the participants in the data collection experiment, but the predicted ones (Figure 5.2, bottom). To do this, we use the model trained in Section 5.3.2, and feed as input the predicted psychological characteristics from the Random Forest model in Section 5.3.2.

The model achieves a mean absolute error of 1.37 (Table 5.2). As expected, there is a drop compared to the 0.98 error that we got using as input the true psychological characteristics. Nevertheless, we notice that the prediction error is not significantly worse than the results reported in Chapter 4, despite using predicted values as input (**RQ2**). This confirms the predictive potential of the psychological characteristics of situations. However, it also suggests the need for more research towards predicting these psychological characteristics more accurately, since that would lead to an overall better prediction of the priority of social situations.

5

Table 5.2: Mean Absolute Errors of the models in predicting the priority of social situations when using different inputs. Results marked with * are significantly different from the others (Wilcoxon Rank sum test, $p < 0.05$).

Model input	Mean Absolute Error in Priority Prediction
Social situation features [96]	1.35
True psychological characteristics of situations	0.98*
Predicted psychological characteristics of situations	1.37

5.4. Study 2 - Evaluating Explanations

In this section we present the setup of the user study we performed to evaluate explanations given by a hypothetical personal assistant agent about why they suggest attending a specific meeting, based on Level 1 and Level 2 information (**RQ3** and **RQ4**).

In this study[5], subjects were presented with pairs of social situations (in this case, meetings), and suggestions from a personal assistant agent regarding which meeting to

[5]The survey questions and the data can be accessed in the supplementary materials in https://doi.org/10.4121/16803889.

attend, followed by an explanation that included as a reason either Level 1 or Level 2 information. Subjects were asked to evaluate these explanations (Figure 5.3). The results of this study are presented in the next section.

5.4.1. DESIGN CHOICES AND MATERIAL

In this section we present the choices we made in the design of the experiment, and the resulting material used for conducting it.

SIMPLIFICATIONS

This study falls under the human grounded evaluation category proposed by Doshi-Velez and Kim [40]: a study with real humans, and a simplified task. The first simplification we made had to do with the fact that subjects were presented with hypothetical scenarios and explanations. This simplification was necessary since we do not yet have a fully fledged support agent ready to use and be tested in practice. Since the proposed scenarios were provided by us rather than by the participants themselves, this comes with the risk that participants may not actually encounter that particular situation themselves in their own lives directly (e.g., some scenarios refer to meetings with work colleagues, however the participant might not be employed). For this reason, in this study we opted for a third-person perspective, i.e., asking participants to imagine how another user might evaluate the explanation if they were to encounter that scenario. Moreover, using existing scenarios allowed us to balance which psychological characteristics were used, which was important for investigating whether people hold different preferences for different characteristics. The second simplification had to do with the fact that the explanations were not formed using a specific explainable AI method, but designed by the researchers based on insights from our predictive models in Section 5.3.2.

In order to make the hypothetical setting as realistic as possible, scenarios were retrieved from the the data collected Chapter 6. In that study, subjects described social situations from their lives, and answered questions about the psychological characteristics of those situations (Level 2). However, the dataset did not include annotated Level 1 information, which is needed to form the explanations based on this type of information. To perform the annotation, we used information that is available in the description of the situations. For instance, if the description says ‘I am meeting my boss to discuss the project’, we infer that the role of the other person is *supervisor*, the hierarchy level is *higher* and the setting is *work*, and consider the information that is not available in the description to be equal across situations. Using only explicit information available in the description to infer Level 1 information allows this procedure to be unambiguous. At this point, we have a dataset with situations described by people, annotated in terms of their social situation features and psychological characteristics which will be used to form the explanations.

SELECTING WHICH INFORMATION IS INCLUDED IN EXPLANATIONS

For an explanation to be realistic, it needs to be based on information that contributed to the suggestion of the agent. In order to find the Level 1 and Level 2 information that is more likely to have contributed to the priority prediction, we identified the features that have the highest weight when predicting the priority of social situations using the Tree-Explainer method of the SHAP package [106]. For Level 1, these features were setting, help dynamic, role, relationship quality, age difference, and shared interests. For Level 2, these features were duty, intellect, positivity and negativity. We assume that the best explanation can be found in this pool of features, since they are the best predictors of priority.

Selecting scenarios

We want users to evaluate the type of information included in the explanations, rather than evaluate whether the agent selected the right feature to include in the explanation. To facilitate this, we formed pairs of scenarios in such a way that both meetings have a set of common situation features/psychological characteristics and a single differing one, which would then be used in the explanation. This was done using the following procedure:

- *Level 1* - Each meeting is annotated with a set of social situation features. To form pairs, we selected scenarios that have the same amount of information in terms of social relationship features (i.e., same number of social situation features known), and that differ in only one social relationship feature.
- *Level 2* - Each meeting is annotated in terms of its psychological characteristics, rated on a scale from 1 (very uncharacteristic of the situation) to 7 (very characteristic of the situation). We consider psychological characteristics with a score higher than 4 to have a *high relevance* in the situation, and those with a score lower than 4 to have *low relevance*. To form pairs, we selected scenarios that have a similar level of relevance (i.e., either high or low) for all psychological characteristics except for one, which has a differing level of relevance.

In total we formed eight pairs of scenarios, where the differing social relationship features were setting, help dynamic, role, relationship quality, age difference, and shared interests. The differing psychological characteristics were duty, intellect, positivity and negativity (two pairs for each). For instance, one of the pairs was:

Meeting 1 - Alice has planned to meet a colleague because they want to update each other about their work.
Meeting 2 - Alice has planned to meet another colleague because the colleague needs her help to solve a work task.

In this case the differing social relationship feature was the help dynamic[6], which was *neither giving nor receiving help* for the first meeting and *giving help* in the second (as inferred from the scenario descriptions), whereas the differing psychological characteristic is the level of duty, which was higher in the second meeting (as annotated by the subjects who proposed these scenarios).

Selecting Agent Suggestions

To determine which meeting the agent should suggest the user to attend, we used a heuristic procedure based on the prediction models from Section 5.3.2. Through the

[6]The feature *help dynamic* can take the values *giving help, receiving help, neither giving nor receiving help.*

5

TreeExplainer method [106] we determined whether each differing feature contributes to a higher or a lower priority level. Since meetings differ in one feature (for each of Level 1 and Level 2), that feature is used as the tie breaker to determine which scenario should have higher priority. Scenarios were selected in such a way that the agent would make the same suggestion regardless whether it uses Level 1 information or Level 2 information for the prediction. This was done to minimize the effect that the agent suggestion has on the evaluation that the subjects give about the explanations. For the aforementioned pair, Meeting 2 has a higher priority because, based on the prediction models:

- Meetings where someone is expected to give help have a higher priority (Level 1 information);
- Meetings with a higher level of duty have a higher priority (Level 2 information).

SELECTING EXPLANATIONS

To form the explanations, we followed insights from research on Explainable AI which suggests using shorter explanations that have a comparative nature [115], [169]. For this reason, explanations include only the differing feature between the meetings (one for each explanation), and are phrased as comparisons between the available choices. For the previously introduced pair of scenarios, the explanations would be:

Explanation based on Level 1 information - Alice should attend Meeting 2 because she is expected to give help, while in Meeting 1 she isn't, and meetings where one is expected to give help are usually prioritized.

Explanation based on Level 2 information - Alice should attend Meeting 2 because because it involves a higher level of duty, which means she is counted on to do something, and meetings involving a higher level of duty are usually prioritized.

5.4.2. MEASUREMENT

In order to evaluate how good the explanations are, we first need to decide on a set of criteria based on which they can be evaluated. Vasilyeva et al. [168] suggest that the goal of the explainer is key in how the explanations are evaluated. Different goals of explainable systems identified in the literature are transparency, scrutability, trust, persuasiveness, effectiveness, education, satisfaction, efficiency and debugging [29], [162], [170]. In our setting, the objective of the personal assistant agent is to justify its suggestions so the user can decide to accept them or not. Therefore, its main goal is to offer clear and understandable explanations for the reasons behind the suggestion, which relate to the goals transparency and satisfaction. Furthermore, we want to assess the persuasive power of the explanations.

To assess how clear the explanations are, we use an adapted version of the explanation satisfaction scale [73]. From the scale, we use the following statements:

- The explanation of [...] is *satisfying*;
- The explanation of [...] has *sufficient detail*;
- The explanation of [...] seems *complete*;

We do not include the items of the scale that refer to accuracy, trust, usefulness to goals and whether the explanation tells the user how to use the system, since these items are not related to the goals of the envisioned support agent.

To further inquire about the clarity and understandability of the explanations, we add the following statement:

- The explanation of [...] is in line with what you consider when making similar decisions;

This is done because we expect that being presented with information which is similar to what they consider when making similar decisions would make the explanations more understandable for the user.

Lastly, another goal of the agent is persuasiveness, which means how likely are the explanations to convince the user to follow the suggestion. This was captured through the following question:

- The explanation of [...] is likely to convince Alice to accept the suggestion.

These items were rated on 5-points scales which were different for each experimental setting, as specified in Section 5.4.4 and Section 5.4.4.

5.4.3. Participants

In total, we recruited 290 subjects through the crowd-sourcing platform Prolific Academic. Participation was open to members that had listed English as their first language. Every subject was compensated for the time they spent completing the study, as per the guidelines of the platform. The study consisted of two experiments. For the first experiment we recruited 100 subjects. Of these, 55 were female, and 45 were male, with a mean age of 31.1 and a standard deviation of 11.8. For the second experiment we recruited 190 subjects. Of these, 108 were female, 80 were male, 1 selected the option 'other', and 1 selected the option 'prefer not to say'. They had a mean age of 29.98 with a standard deviation of 10.28.

5.4.4. Procedure

In this section we introduce the procedure that was used for this study. The study consisted of two experiments. In the first experiment (between-subject design, **RQ3**, top part of Figure 5.3), participants are shown either an explanation based on social situation features (Level 1 information), psychological characteristics of the situation (Level 2 information), or a control explanation based on features that were considered not useful. In the second experiment (within-subject design, **RQ4**, bottom part of Figure 5.3), we show participants both Level 1 and Level 2 explanations for a specific suggestion by the agent, and ask them to *compare* these explanations and indicate which one they prefer. Both experiments were conducted as online surveys, and the subjects were recruited through the crowd-sourcing platform Prolific Academic. The study received the approval of the ethics committee of the university. The experimental procedure was similar in both experiments:

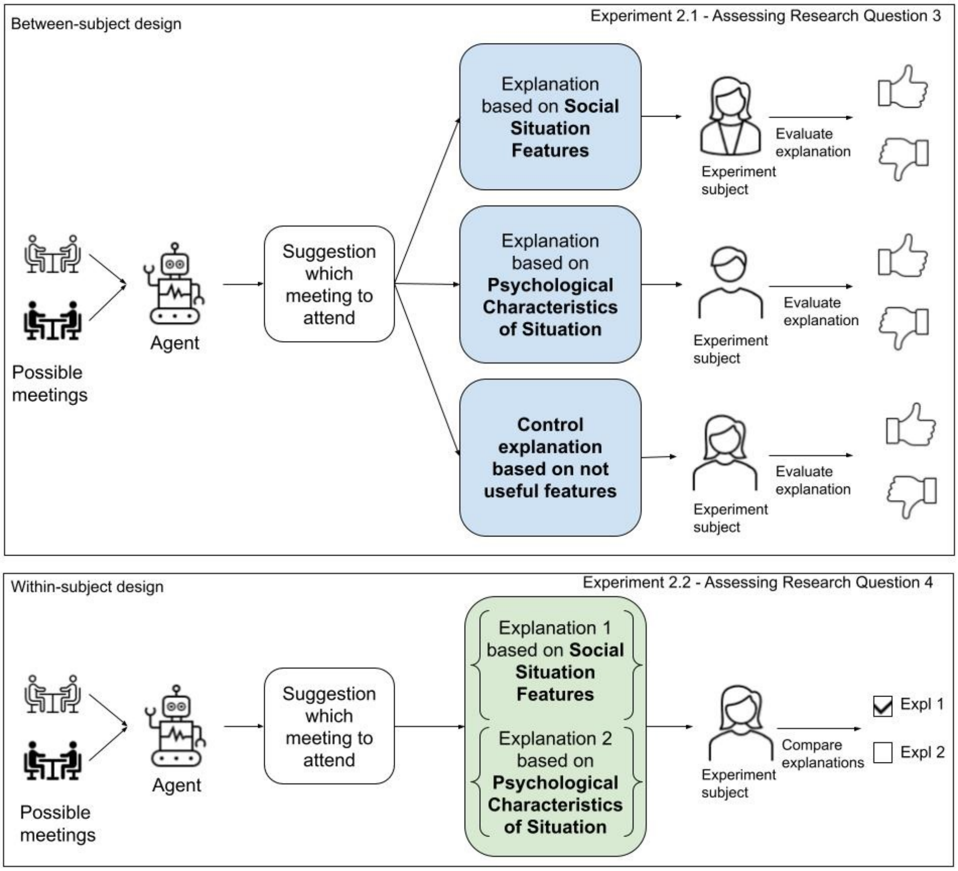

Figure 5.3: Conceptualization of Study 2, used to assess Research Questions 3 (top part) and 4 (bottom part).

- *Introduction* - Subjects were informed about the study and were presented with the consent form.
- *Demographics* - Subjects were asked about their age and gender to check whether the population sample was sufficiently broad.
- *Case-study* - Subjects were introduced to Alice, a hypothetical user of the socially aware personal assistant agent. Subjects were told that during a specific week Alice is particularly busy, so the agent makes suggestions which meetings she should attend and which ones she should cancel.
- *Scenarios* - Subjects were presented with a pair of meeting scenarios, and they were asked which meeting they would suggest Alice to attend. This was asked to control for biases that they would have regarding the agent's suggestions, in case

their own opinion differed from that of the agent. Furthermore, in an open question they were asked about the reasons behind this suggestion. This was asked to get more insights into the reasoning process of subjects in such situations. In total subjects were presented with four pairs of scenarios.

- *Evaluation of explanations* - Subjects that made suggestions in line with the agent were presented with the full questionnaire which included all measures from Section 5.4.2. Subjects that made suggestions that were different from what the agent would suggest were presented with a question regarding the persuasiveness of the different explanations (namely: "The explanation offers convincing arguments"). This was done to take into account biases: We expect that subjects that do not agree with the agent suggestion would be implicitly evaluating the suggestion rather than its explanation.

In the next subsections we present the specifics of each experiment.

EXPERIMENT 2.1

This part of the study had a between-subjects design. Subjects were presented either with explanations based on Level 1 information, Level 2 information, or they were part of the control group, which we added to serve as a baseline. In related work (e.g., [169]), control groups normally do not include an explanation, since the goal is usually to evaluate the impact of the explanation in the overall quality of the suggestion. However, in our setting that would be obsolete since the questions specifically refer to explanations. For this reason, in the control group subjects were presented with explanations that included information that could in principle be useful for determining the priority of meetings, but did not make sense for those specific scenarios. Explanations in the control group included information such as weather, geographical location or time. For instance, an explanation was "Alice should attend the first meeting because it is spring".

This design presents subjects with only one type of explanation, so the evaluation is absolute rather than relative to the other explanation types. This allows us to answer **RQ3**: to what extent can social situation features and psychological characteristics of situations be used as a basis for explanations?

The aforementioned measurements were presented as statements such as "The explanation provided about the reasons why the agent suggests Meeting 2 is satisfying". Subjects could answer on a 5-point Likert scale, ranging from Strongly disagree to Strongly agree.

EXPERIMENT 2.2

This part of the study had a comparative within-subject design. This design presents subjects with two explanations for each pair of scenarios: one based on Level 1 information, and one based on Level 2 information. Through this setting, we address **RQ4**: when do people prefer one type of explanation versus the other? The measurements were framed as comparisons, for instance "Which explanation do you consider more satisfying?". Subjects could answer 'Significantly more Explanation A', 'Slightly more Explanation A', 'Both equally', 'Slightly more Explanation B' and 'Significantly more Explanation B'.

5.4.5. RESULTS AND DISCUSSION

In this section we present the quantitative results of the two user studies described above, and we analyze the answers to the open question.

EXPERIMENT 2.1

Each of the subjects was presented with four pairs of scenarios, which means 400 pairs of scenarios were shown to subjects across the different conditions (128 pairs in the Level 1 group, 140 pairs in the Level 2 group, and 132 pairs in the control group). In 73% of the total cases, subjects would suggest Alice to attend the same meeting that the agent would suggest. Figure 5.4 presents the subjects' answers for each of the measurements regarding the explanation provided by the agent. This applies to the subjects whose suggestions were in line with the suggestions of the agent.

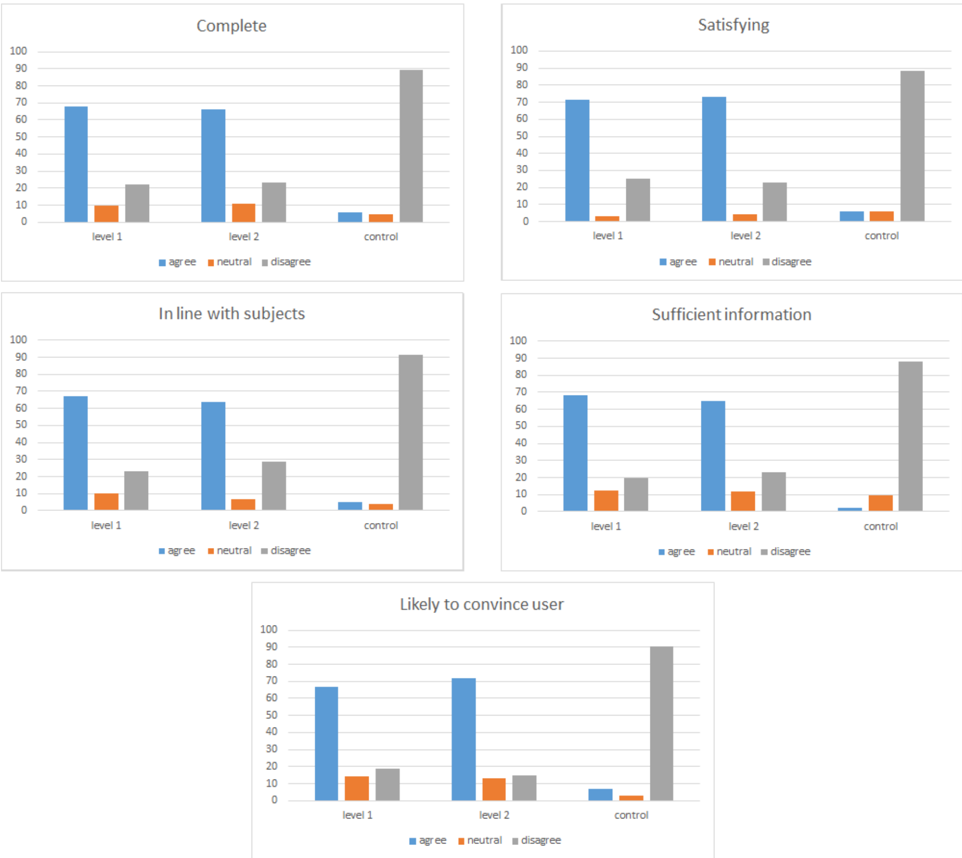

Figure 5.4: Answer distributions for the different measurements. The x axis represents the answer options for each of the levels. 'Strongly agree' and 'Somewhat agree' were grouped together as 'agree', and 'Strongly disagree' and 'Somewhat disagree' were grouped together as 'disagree'. The y axis shows the percentage of subjects that gave a specific answer.

The majority of the subjects considered the explanations based on Level 1 or Level 2 information to be complete, satisfying, in line with how the subjects reason, likely to convince the user, and having sufficient information. While explanations based on Level 1 or Level 2 information were thus considered positively, on the other hand, subjects strongly disliked the explanations offered in the control setting. This confirms that the positive effect was not just due to the presence of an explanation as such, since subjects do not give a positive evaluation to an explanation which does not apply to the suggestion.

The answers of the subjects whose suggestions were not in line with the suggestion of the agent are presented in Figure 5.5. We see that subjects do not find the explanations of the agent to provide convincing arguments. This shows that there is some inherent bias, and that subjects are implicitly evaluating the quality of the suggestion too, and not just the explanations. However, we notice that explanations containing Level 2 information are still seen as convincing in 40% of the cases, compared to 21.6% for explanations containing Level 1 information.

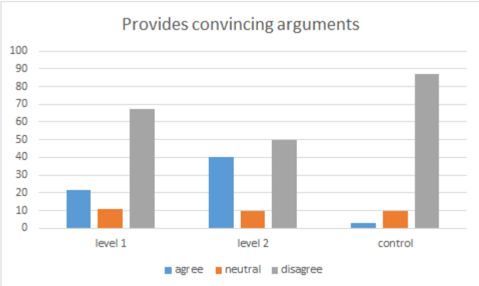

Figure 5.5: Answer distribution for the subjects who would make a suggestion different from the agent's.

To control for statistical significance we perform the Kruskal-Wallis test, a non-parametric version of ANOVA which can be applied to non-normally distributed data like in our case. Results showed that there is significant difference between the condition means for each of the measurements ($p < 0.001$). To control for differences between the pairwise conditions, we perform Dunn's test. Results show that the evaluation of both level 1 and level 2 explanations are significantly different from the explanations of the control group across all measurements ($p < 0.01$). However, when comparing the evaluations of level 1 explanations to those of level 2 explanations, the difference is not statistically significant for any of the measurements ($p > 0.05$).

This experiment allows us to answer **RQ3**: Approximately 70% of the subjects find the explanations based on Level 1 or Level 2 information to be complete, satisfying, in line with the way the subjects reason, likely to convince the user, as well as containing sufficient information. This makes such information a good candidate for forming explanations in personal assistant agents.

Experiment 2.2

The goal of Experiment 2.2 was to evaluate **RQ4**. Results are presented in Table 5.3. First of all, for each measurement we report the answer distributions across the different sce-

Table 5.3: Significance test could not be performed for the measurement ‘provides convincing arguments’ since only a small portion of subjects made choices different from the ones of the agent and was presented with that measurement.

	Preferred Explanation	Satisfying	Sufficient information	Complete	In line with user	Likely to convince	Convincing arguments
Duty-salient situations	Level 1	36.2%	22.9%	31.3%	38.6%	30.1%	40.1%
	Neutral	7.2%	19.3%	15.7%	12%	12.1%	18.2%
	Level 2	56.6%	57.8%	53%	49.4%	57.8%	40.1%
Intellect-salient situations	Level 1	43.4%	34.9%	44.6%	45.7%	47%	50%
	Neutral	20.5%	26.5%	24.1%	14.5%	12%	25%
	Level 2	36.1%	38.6%	31.3%	39.8%	41%	25%
Negativity-salient situations	Level 1	59%	57.8%	59%	53%	57.8%	50.6%
	Neutral	14.5%	19.3%	22.9%	14.5%	14.5%	20.3%
	Level 2	26.5%	22.9%	18.1%	32.5%	27.7%	29.1%
Positivity-salient situations	Level 1	37.3%	37.3%	37.3%	43.3%	42.4%	18.8%
	Neutral	14.5%	28.9%	26.5%	15.7%	9.6%	18.8%
	Level 2	48.2%	33.4%	36.2%	41%	48.2%	62.4%
Friedman's test	χ^2	19.935	26.417	21.549	4.9594	19.094	-
	df	3	3	3	3	3	-
	p-value	**<0.001**	**<0.001**	**<0.001**	0.17	**<0.001**	-
Post-hoc analysis Conover's test	Duty-Intellect	0.18	0.067	0.07	0.91	**0.02**	-
	Duty-Negativity	**<0.001**	**<0.001**	**<0.001**	0.18	**<0.001**	-
	Duty-Positivity	1.00	**0.02**	0.71	1.00	0.626	-
	Intellect-Negativity	0.488	0.067	0.29	1.00	1.00	-
(p-values)	Intellect-Positivity	0.393	1.00	1.00	1.00	1.00	-
	Negativity-Positivity	**<0.01**	0.199	**0.02**	1.00	0.068	-

nario pairs based on which psychological characteristic was salient in the pairs. The results show that the preferences of the subjects vary between situation types. However, we notice consistency within types: for a specific pair, subjects tend to prefer the same explanation across all measurements. Given this, for simplicity we will abuse terminology and say that subjects prefer one explanation over the other in a pair of scenarios when the subjects prefer that explanation for at least four measurements.

From the answer distributions, we notice that in situations where duty is the salient feature, subjects prefer explanations involving Level 2 information. On the other hand, in situations where negativity is the salient feature, subjects strongly prefer explanations involving Level 1 information. This seems to suggest that subjects do not like explanations that have a negative framing[7]. For situations where the salient feature is intellect or positivity we cannot reach a clear conclusion regarding which explanation is preferred, since the results are different across pairs and seem to be context dependent.

To control for statistical significance we perform Friedman's test, a nonparametric alternative to repeated measures ANOVA since our data is measured on an ordinal scale rather than continuous. For each measurement, the test controls whether the answers in each situation type (Duty-salient, Intellect-salient, Negativity-salient and Positivity-salient) differ. Results show that the answer distributions significantly differ ($p < 0.05$) for all measurements apart from ‘in line with subject’. Friedman test is an omnibus test

[7] The explanation involving Level 1 information was *“Alice should attend Meeting 2, since in it she is meeting someone with whom she has a better relationship, and meeting with people with whom one has a better relationship are usually prioritized."*, while the explanation involving Level 2 information was *“Alice should attend Meeting 2, since Meeting 1 could entail a high level of stress, and meetings that entail a low level of stress are usually prioritized."*

statistic, which indicates that there are significant differences in which explanations are seen as more satisfying, complete, having more sufficient information and likely to convince the user based on situation type, but does not tell which specific situation types have a significant effect on these measurements. For this, we conduct a post-hoc analysis in which we performed the Conover's test for pairwise comparisons in situation types. Confirming the insights from the answer distributions, we notice that the preferred explanations in situations where Duty is the salient feature significantly differ from situations in which Negativity is the salient feature. For the other situation types there is no significant effect across measurements.

This experiment gives some insights towards answering **RQ4**. It shows that subjects prefer explanations involving Level 2 information when duty is the salient feature, and explanations involving Level 1 information when negativity is the salient feature. However, this experiment also shows that more research is needed to determine which type of explanation is preferred for each situation. Overall, an agent that can give explanations including information from either level is beneficial, since the preferred explanation is context dependent and can vary.

OPEN QUESTION ANALYSIS

After answering which meeting they would suggest to Alice, subjects were also asked about the reasons behind this suggestion. This was done to assess the type of information that users would include in their reasoning, and how it compares to the explanations given by the agent. The results are presented in Figure 5.6. The answers were analyzed by the first author in a two step procedure, following guidelines from Hsieh and Shannon [75]. The first step involved summative content analysis. In it, each open answer was labeled to refer to Level 1 information, Level 2 information, or neither. To assign a label, keywords for Level 1 information were extracted from the social situation features, whereas keywords for Level 2 were extracted from the descriptors of the psychological characteristics of situations. The second step involved the open answers which did not fall under Level 1 or Level 2 information. For these answers, we performed conventional content analysis. This involves coming up with categories based on the data, rather than using preconceived categories. After reading the answers multiple times, keywords were highlighted as labels, and then clustered in cases when the keywords are logically connected. This analysis is exploratory and does not intend to provide comprehensive answers on the reasons that users have for deciding between meetings.

The results show that in more than half of the cases, subjects offered a reason that involved either the Level 1 or the Level 2 relevant feature for that pair. This confirms that subjects also reason themselves in terms of this information in many cases. Level 1 information was mentioned significantly more than Level 2 information, but this was to be expected since Level 1 information is directly present in the description of the meetings, so it is more salient.

From this open question we can also extract other types of information that users find relevant. For instance, in 12% of the cases subjects gave a reason that was related to temporal aspects, such as 'Meeting 1 is more urgent', or 'Meeting 2 is more difficult to reschedule'. This feature should be considered for inclusion to the list of Level 1 situation features, since it was consistently mentioned by subjects. Two other reasons that

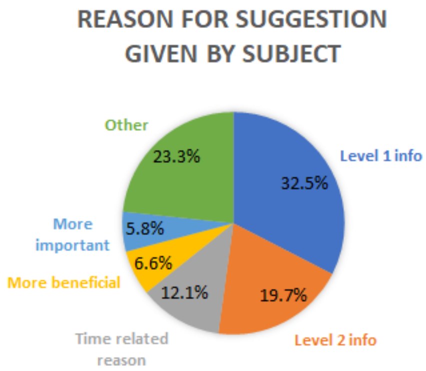

Figure 5.6: Distribution of reasons given by the subjects when asked why they would suggest attending a specific meeting.

5

were consistently mentioned were ‘more beneficial’ and ‘more important’. Subjects also mentioned various other similarly vague answers (e.g. ‘better’) which did not appear consistently, therefore were clustered under ‘other’. Such answers show that subjects often do not explicitly dig deeper into the reasons, but offer only superficial ones.

When taking a closer look at subjects who in the open question used Level 1 or Level 2 information, we notice that the reasons that the subjects give do not necessarily match with their preferred explanations. In 43% of the cases, in the open question subjects gave as a reason for their suggestion information from one of the levels, and in the questionnaire they preferred the explanation that included information from the other level. For instance, in the open question for Pair 5 one of the subjects says *“Meeting two will be more enjoyable and less stressful"*, which fits almost perfectly with the explanation given by the agent that involves Level 2 information. However, in the questionnaire this subject always prefers significantly more the explanation that includes Level 1 information. This ‘flip’ happens in both directions: in 50% of cases it’s from Level 1 to Level 2 and in 50% the other way around. This suggests that there are users that want to hear explanations that differ from the reasons that they thought about themselves, providing another perspective on which explanations the agent should provide to the user.

5.5. CONCLUSION

5.5.1. DISCUSSION

In this work, we explore the effect of incorporating information about the psychological characteristics of situations in a socially aware agenda management support agent. To assess the benefits of this approach, we evaluate its contributions in improving the accuracy of the agent predictions, as well as in providing more satisfying explanations for the suggestions to the user.

Automatic agenda management has been previously used as a test bed for studying how to model social relationships in agents. For instance, Rist and Schmitt [141] introduce an agent that negotiates meetings on behalf of the user. The agent incorporates in its negotiation process information regarding how important the meeting is for the user, as well as information regarding the relationship quality between the user and the other person. Such an agent would benefit from the ability to automatically assess the priority of the different meetings from the point of view of the user. We hypothesized that the priority of meetings can be accurately predicted using as input the psychological characteristics of the meeting. Results in Section 5.3 show that psychological characteristics of situations are a significantly better predictor of the priority of situations than social situation features, thus supporting our hypothesis. Thus, using our approach for predicting the priority of social situations would be beneficial for support agents. Asking the user about the psychological characteristics of each individual situation would be a cumbersome task. For this reason, we explore whether this information can be assessed automatically. We show that using a random forest model that take as input the social situation features of a situation allows us to accurately predict the psychological characteristics of that situation. Collecting social situation features is a less invasive task, since information about social relationships can be collected once and used across multiple interactions. Murukannaiah et al. [118] show that active learning can be used to collect information in a less invasive manner.

In Section 5.4, we show that people find explanations based on social situation features and psychological characteristics of situations to be satisfying, containing sufficient information, complete, in line with how they think, and convincing. Using brief explanations focusing on why a certain suggestion was made as opposed to the alternative led to satisfying explanations, in line with findings from related work [115], [138]. Furthermore, we notice that when the suggestions of the agent are not in line with people's expected suggestions, they do not like the explanations. This is in line with findings reported by Riveiro and Thill [142]. Work on explanations for recommender systems [163] suggests that the type of information contained the explanation affects the perceived quality of the explanation. Our work represents a first attempt in evaluating what type of information is preferred in recommendations regarding social situations. Our findings show that people prefer explanations based on psychological characteristics in situations where the level of duty is relevant, and explanations based on social situation features in situations where the level of negativity is relevant. Both types of explanations were evaluated positively, indicating that it may be beneficial if support agents were able to give explanations based on both types of information.

Overall, our results suggest that incorporating information about psychological characteristics of the user's situation can be beneficial for support agents, since it would enable them to more accurately predict information that can be used as a basis for suggestions and for explaining the suggestions to the user.

5.5.2. ETHICAL IMPACT

Several ethical considerations have to be made before deploying an agent to offer support in the real world. First of all, the agent's assessments of the priority of situations can be inaccurate, thus offering to the user suggestions that can have social repercussions. For this reason, in our use case the decision remains in the hands of the user, and the agent also offers explanations for its suggestions. However, this also does not fully mitigate ethical risks. For instance, the agent might wrongly infer that a specific social situation has a high level of negativity, and inform the user about it in an explanation. However, if this is a situation which is sensitive for the user, the explanation can cause distress. Therefore, it is important to increase prediction accuracy, as well as to have more studies that assess the effects on a user of using such an agent on a daily basis.

5.5.3. LIMITATIONS AND FUTURE WORK

In this work, results were based on the use case of a socially aware personal assistant agent. Future work should extend the findings for different types of support agents and other support domains. Here it will be particularly interesting to investigate if the general nature of psychological characteristics makes them a good candidate to predict other aspects of social situations besides their priority. Assuming a support agent that can assist in various tasks and different daily situations, having a common conceptual grounding for assessing the meaning of situations for the user could have advantages for human-machine meaning. Furthermore, in this chapter we used a hypothetical setting in order to be able to gather larger amounts of data in a controlled way. Based on the results from this hypothetical setting, it is important to build a prototype support agent in order to test the methods in real tasks.

While answering Research Questions 1 and 2 we found that predicting the psychological characteristics of situations accurately is crucial in order to better predict the priority of situations. In future work, we will explore other techniques, such as using natural language processing techniques to extract the psychological characteristics of situations from textual descriptions of situations. Lastly, Study 2 shows that while both social situation features and psychological characteristics of situations can be the basis of explanations given by support agents, more research is needed to determine which type of explanation to give in which situation.

6

Assessing how Social Situations Affect Personal Values

Support agents are investigated more and more as a way of assisting people in carrying out daily tasks. Support agents should be flexible in adapting their support to what their user needs. Research suggests that the situation someone is in affects their behaviour, however its effect has not been incorporated in the decision making of support agents. Modelling the characteristics of situations explicitly and studying their effect on internal perceptions of the user, such as their personal values, would enable support agents to provide more personalized support. We propose a method which groups situations according to their psychological characteristics, and in turn determines which personal values of the user would be promoted or demoted in each group of situations. To do this, we conduct a user study to gather data from participants about situations that they encounter in their daily lives. Results show that the created groups of situations significantly promote or demote certain personal values. This approach can allow support agents to help the user in a way which is in line with their personal values.

This chapter was published in the Proceedings of the International Workshop on Modelling and Reasoning in Context (MRC2020) [94].

6.1. Introduction

Kurt Lewin, already 80 years ago, proposed that human behaviour is a function of both the personality of the person, as well as the situation in which they are in [101]. This is now a widely accepted idea in social psychology, after multiple debates in the field [135]. However, applications of support agents (e.g. [82], [120], [160]) focus mostly on modelling internal aspects of the user. Personal values are one of these aspects. They represent what is important to people [53], and because of that, they guide behavior. However, how important a certain value is for the user is not the only factor that guides behaviour. Whether that value is actually relevant in a given situation also plays an important role. For example, the fact that having an exciting life is important to someone plays a role in deciding the next holiday destination, but most probably does not affect the decision whether to have pizza or salad for dinner. On the other hand, the fact that someone values health would affect that decision, since having salad is supposed to promote the value health (i.e., help you fulfill it), whereas having pizza can demote it (i.e., prevent you from fulfilling it). This means that apart from personal values, it is important to also consider how the situation in which someone is in affects those values. This information can be used by a support agent in combination with information about the value preferences of the user in order to offer support on how to handle daily life situations. Continuing the previous example, the agent would suggest having a salad to a user that finds health important.

In this chapter we explore the relationship between the situation in which a user is in, and the personal values that are affected by the situation. To achieve this, first of all we explore ways how to group similar situations together. To do so, we extend the work on Context Space Theory [123], which refers to a group of similar situations as a *subspace*. A situation subspace is a group of situations which have the same range of numerical values on certain attributes (Section 6.3.1). In this work, we use psychological characteristics of situations as attributes. Psychological characteristics are seen as dimensions that can be used to describe situations, similar to the manner in which people can be described with traits, attributes, or qualities [44]. Examples of these characteristics are positivity, duty, intellect, mating etc [135] (Section 6.3.2). This leads to the following research question:

- What methods can we use to group situations according to their psychological characteristics as context attributes?

Then, we investigate whether the identified subspaces significantly promote or demote personal values. Our research hypothesis is:

- Situations of the same subspace significantly promote or demote the same personal values, in comparison to a general population of situations.

While the research question and hypothesis guide the work presented in this chapter, we do not aim to provide definitive answers here. Rather, as this is a novel research direction, our aim is to assess the feasibility of the approach as a basis for future work, as we proposed in previous work [91]. Our results indicate that it is possible to group situations into subspaces by using domain knowledge and insights from the data, and that situations from the same subspace tend to promote and demote the same personal values.

The rest of this chapter is structured as follows: In Section 6.2 we present a high level architecture of our approach, and compare it to related work. In Section 6.3 we motivate our research choices for the use of psychological characteristics to group situations into subspaces, and we provide a short introduction to the concept of personal values. In Section 6.4 we present the user study in which we gather data in order to build the method which we described in the architecture. We present and discuss the results in Section 6.5, showing that situation subspaces can promote or demote specific personal values. Section 6.6 concludes this chapter.

6.2. AGENT ARCHITECTURE

We propose an architecture which explains how a support agent can use information about the psychological characteristics of situations in order to determine the promoted or demoted personal values, and in turn offer support to the user. The architecture (Figure 6.1) depicts two main components: a learning component in which we use data gathered from people to identify situation subspaces, and a support agent which uses this information to provide support to the user.

In the first component, participants of a user study describe situations from their lives and provide us with the psychological characteristics as well as the promoted and demoted values of these situations (Section 6.4). We use these psychological characteristics together with domain knowledge in order to form situation subspaces (Section 6.5.2). Then, we determine the promoted or demoted values for each situation subspace (Section 6.5.3). When the support agent is interacting with the user, once presented with a new situation, the agent uses the subspace rules to classify the situation to a subspace, as done in Context Space Theory [123]. By knowing the subspace values, the agent can reason about the promoted or demoted values of that specific situation. This information, in combination with the value preferences of the user, can be used in order to reason about support. This last part is not tackled in this chapter, but is displayed in the architecture in order to make the bigger picture clear.

This approach would allow support agents to align their suggestions with the personal values of their users. Let us consider an agent that recommends free time activities to the user, and the options are going to a party and attending a workshop to learn a new skill. Following the architecture depicted in Figure 6.1, the agent might infer that the first would promote pleasure, and the second would promote capability. This way, the agent would suggest going to a party to a user who prefers the value pleasure, and attending the workshop to a user who prefers the value capability.

Related work Other work also focuses on using concepts such as personal values and context in socio-technical systems, in order to enable them to understand and adapt to human motivations. We introduce some of these approaches in order to position our work. Tielman et al [161] propose an approach to derive norms from a combination of values, context and actions. Context is used as a modifier to determine how much a value is promoted or demoted when performing a certain action, and this information

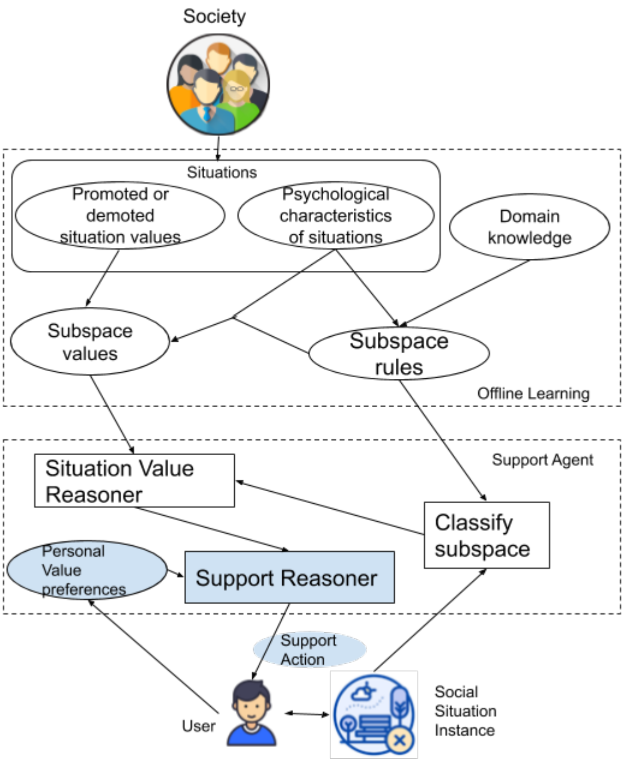

Figure 6.1: High-level architecture of the approach. Concepts shaded in blue represent aspects which we do not directly tackle in this work (i.e. Modeling the user preferences, extracting psychological characteristics of situations, and reasoning about the type of support). Circles represent knowledge elements (e.g. personal value scores, subspace rules), whereas squares represent reasoning steps. Arrows indicate the workflow of the approach. Icons used in the architecture were made by Freepik and retrieved from www.flaticon.com

is elicited from the user. Context is not modelled explicitly, and can be represented by a list of variables, depending on the situation. Similarly, Cranefield et al. [31] propose an approach on how to use values in order to help users with moral decisions. The work focuses on the reasoning about aligning the values of the user to the values that are promoted or demoted by different actions. Similarly, the values and context are assumed to be predetermined. Our work focuses on the other point of view: how to actually infer what values are promoted or demoted in a given situation? In a way, our work can be considered an extension of these approaches, since the output of our work can be used as an input for these reasoning frameworks. Kayal et al [82] also take a step in this direction. In their work, they ask participants about their personal values and about the promoted and demoted values of different social commitments. They then use this information to break ties when different commitments overlap. Other work (e.g. [87], [99]) describes the relation between the environment and the people in terms of contextual affordances, which represent potential actions that the environment (or parts of it) allows people to perform. For instance, a chair allows the action “sit". This is in principle similar to what we are doing, since personal values can be seen as affordances of a situation, since some situations allow people fulfill specific personal values. For example, a situation in which a person is exercising would help them fulfill the personal value of health.

6.3. SITUATIONS AND PERSONAL VALUES

6.3.1. SITUATION SUBSPACES

Research in computer science uses terms such as situation awareness (e.g. [47]) and context awareness (e.g. [4]) to describe attempts to enable artificial agents to better understand their surrounding environment. According to Barwise [11], these concepts refer to the same thing, and situations represent a way of modelling contexts. Other researchers (e.g. [7]) see context as a lower level of abstraction, and situations can be seen as “logically aggregated pieces of context". In Endsley’s situation awareness framework [47], the aforementioned interpretation of context would refer to the situation cues in the perception level of situation awareness. There is vast research on modelling and reasoning about context and situations, and describing this research in depth is beyond the purview of this chapter. For a detailed account, readers can refer to existing surveys [17], [175]. In this section, we introduce possible approaches on how to use context elements in order to determine the promoted and demoted values of a situation.

Our proposed approach is to first group similar situations into so-called situation subspaces, and then to determine the promoted and demoted values of that subspace. This is inspired by work on Context Space Theory [123]. In their approach, context is represented as an object in a multidimensional Euclidean space, called situation subspace. A context state is represented in terms of attributes, and each dimension of the situation subspace represents an accepted region for a specific context attribute. This way, when given a set of attributes that define a context state, we can infer whether the state is or is not part of the situation subspace. For example, a situation subspace can be “Person is healthy" and its attributes are “Body temperature" with an accepted region of values between 36.0 and 37.5 and “Resting heart rate" with an accepted region of values between 60 and 100. In our approach, we consider a situation subspace to be the set of situations having similar psychological characteristics (Section 6.3.2). For instance, a subspace can consist of situations where the characteristics Duty and Intellect have a value between 4 and 7.

Using situation subspaces facilitates the process of explaining the suggestions of the support agent to the user, since each subspace is defined by a set of attributes. The reasoning is explicit: for instance, situation X is in subspace A because of attributes B and C, and situations in subspace A promote values Y and Z. These steps can be available to the user. Furthermore, this way of approaching situations is also in line with work on social psychology on how people actually deal with situations. Gigerenzer [59] suggests that people have different modules of interaction, and when presented with a new situation they “classify” it as part of one of the modules, and then follow the “interaction script" of that module.

One other option for reasoning about the values of a situation would be to look at the correlation of each individual psychological characteristic of the situation with specific personal values (e.g. as done by [135]). However, this approach does not take into account the possibility that the ways in which characteristics are combined in a situation also affect the values that are promoted or demoted in it. We explore this possible connection in Section 6.5.3. In the current section we simply give an intuition. For instance, situations with a high level of mating can in general affect the value pleasure, however it is the combination with high positivity or high negativity that affects whether the value is promoted or demoted. Furthermore, if we consider each psychological characteristic individually, it is not clear whether the low score of a characteristic indicates that a value is demoted or not affected. For instance, knowing that situations with high intellect promote capability is not enough to determine whether situations with low intellect demote this value or do not affect it. Our approach takes the potential effect that the combination of psychological characteristics have on personal values into account, but does not rely on it: if that effect does not hold, our approach would simply take into account the correlation between individual psychological characteristics and personal values.

Lastly, we can reason about personal values by training a model that takes as input the situation's psychological characteristics, and predicts the score for each value. This way, the model would actually take into account all the psychological characteristics of the situation and their potential interactions. Putting aside the requirement for high amounts of data and the non-trivial task of building such a model, our primary reason for not following this approach is its black box nature. We believe one of the key features of a support agent is its ability to explain its suggestions to the user. Such a comparison, and the potential trade-off between accuracy and explainability, is something that we plan to explore in future work.

6

6.3.2. Psychological Characteristics of Situations

Research in social psychology has explored ways in which situations can be systematically described. Rauthmann et al. [135] discuss three ways in which situational information can be taxonomized: Cues (e.g. persons, places, objects etc.); (psychological) Characteristics (which attributes can be used to describe situations - e.g. positivity, intellect, duty etc.); Classes (which kind of situations are there - e.g. social situations, work situations etc.).

In this work we focus on the use of psychological characteristics. There are several taxonomies of situations on the psychological characteristics level. We choose the DIAMONDS taxonomy since it covers a wide variety of daily life activities and it provides a validated 24-items survey which allows the measurement of the psychological characteristics of situations through online surveys. Horstmann et al. [74] suggest that the dimensions of the existing taxonomies have a high level of similarity when compared across taxonomies, so our choice should not influence the outcome of the work. The DIAMONDS taxonomy describes situations in terms of the following dimensions:

- **Duty** - situations where a job has to be done, minor details are important, and rational thinking is called for;

- **Intellect** - situations that afford an opportunity to demonstrate intellectual capacity;

- **Adversity** - situations where you or someone else are (potentially) being criticized, blamed, or under threat;
- **Mating** - situations where potential romantic partners are present, and physical attractiveness is relevant;
- **Positivity** - playful and enjoyable situations, which are simple and clear-cut;
- **Negativity** - stressful, frustrating, and anxiety-inducing situations;
- **Deception** - situations where someone might be deceitful. These situations may cause feelings of hostility;
- **Sociality** - situations where social interaction is possible, and close personal relationships are present or have the potential to develop.

There are different reasons for using the psychological characteristics of situations in order to group them. First of all, psychological characteristics allow us to assess similarities between situations beyond their physical cues (e.g. where is the situation taking place, how many people are involved). Social psychology (e.g. [25], [44], [130], [154]) suggests that people think about situations by using their psychological characteristics. They create impressions of situations as if they were real, coherent entities. These impressions allow people to better navigate through the world by being able to predict what will happen and coordinate behaviour accordingly. This inherent psychological component of situations makes them difficult to interpret only in terms of physical context. For instance, let us consider a scenario where our user, Alice, is going out with friends. The relevant physical attributes would be the activity (i.e. going out), the location, time etc. A support agent might determine that such situation promotes pleasure. On the other hand, it is also possible that at some point Alice is going out and some people that she dislikes join. In that case, the situation could actually demote the value pleasure. However, from the point of view of physical cues, everything would remain the same and this difference would not be captured. In Chapters 3 and 4 we proposed a set of social cues that can be used to capture such differences, for example the quality of the relationship with the other person or the level of trust. However, despite capturing the psychological component of situations, these social cues remain a low-level description.

Another advantage of focusing on the psychological characteristics is easiness of explainability. This means the support agent can explain its suggestions to the user in a way that is understandable and intuitive to people. To continue the previous scenario, we assume our support agent wants to propose an activity which promotes the value of pleasure to Alice, since this value is important to her. It would be more intuitive for Alice to understand that the situation “going out with friends" promotes pleasure because it has high positivity and low adversity, rather than because it is an activity that takes place after 8pm, at a bar, and a certain amount of people are present.

Focusing on the psychological characteristics of situations allows us to identify similarities in situations that look very different. For instance, a situation in which a parent is helping their child with a school project and a situation in which that same parent has an important work meeting do not have anything in common when it comes to physical cues, however they both potentially involve a high level of duty and intellect, and promote values such as helpfulness and capability. This also brings forward practical considerations from a technical point of view: there can be a very high number of physical cues that can be measured, and what is actually relevant differs from situation to situation. Furthermore, highly general concepts such as “activity" are difficult to model in a way which actually makes them comparable from a situation to another. For these reasons, deciding which elements to model and how to do it is both crucial and non-trivial. Our approach allows us to abstract from the physical context, which results in a low dimensionality of characteristics that are proven to be relevant across daily situations [135].

There is some difference in terminology when comparing Context Space Theory with DIAMONDS. A context state from Context State Theory is simply referred to as a situation in DIAMONDS, and context attributes would be represented by the situation dimensions. In this work, we will use the DIAMONDS terminology.

6.3.3. Personal Values

Values represent key drivers of human decision making (e.g. [143], [150]). Friedman and colleagues [53] define values as “what a person or group of people consider important in life". People hold various values (e.g. wealth, health, independence) with different degrees of importance. The main features of personal values which make them relevant to our work have been explicitly described by Schwartz [152], but are also implicitly present in other work on values. First of all, values refer to desirable goals that motivate action, and they serve as standards to guide the selection of actions, people, or events. This means that (unconsciously) people’s decisions are influenced by values. Secondly, values transcend specific actions and situations. For instance, values such as honesty are important to someone regardless of the activity they are doing or who they are with. Lastly, what puts this all together is the fact that in order for values to influence action not only should they be important to the actor, but they should also be relevant in that specific context. This suggests that if we know which values are likely to be activated in a certain context (or situation) and have information about the value preferences of a user, we can use that information to evaluate how much does a situation promote or demote personal values that are important to the user. It is also important to notice that in this work, we talk about personal values on three different levels:

- Personal values are important to an individual - e.g. Alice values achievement;
- A specific situation can promote or demote personal values to someone in the situation - e.g. Being a speaker at a conference promotes the value achievement for Alice;
- A situation from a certain subspace usually enables promoting or demoting a personal value to someone in the situation - e.g. Being part of situations with high intellect and high duty usually promotes the value achievement for people.

The most prominent models of human values were proposed by Rokeach [143] and Schwartz [150]. These models are universal and domain-independent, making them suitable for our purpose, in which we will deal with a wide range of every day situations. This is different from other approaches where the first step was to find a subset of values which are more applicable to a certain domain, for instance mobile location sharing [82] or music recommendations [110]. In our work we use the model proposed by Schwartz since it offers validated measurement instruments with fewer items than Rokeach, which makes them more applicable to online surveys. Furthermore, it is to be noted that Schwartz builds on the work of Rokeach and other researchers, so there is overlap in their proposed value lists. The Schwartz theory of basic human values [150] recognizes 10 universal value groups, namely: Self-direction, Stimulation, Hedonism, Achievement, Power, Security, Conformity, Tradition, Benevolence and Universalism. Each of these value groups includes more "specific" values, as depicted in Table 6.2.

6.4. USER STUDY

In this user study we gather data[1] for constructing and evaluating our methods. The study consists of three parts: first, participants were asked to describe situations from their daily lives (part 1), then they had to answer questions about the psychological characteristics of the situations (part 2) and finally they had to answer questions about how much the situations promote or demote certain personal values (part 3). The study was approved by the ethics committee of TU Delft.

6

Participants We collected answers from 150 participants recruited in the crowd-sourcing platform Prolific Academic[2]. Using a crowd-sourcing platform allowed us to efficiently obtain a large sample size in a short amount of time. Respondents received a monetary compensation for the time they spent, as per the platform policies. The average age of participants was 32.38 (SD=12.1). 51.3% were female, 44% male, and 4.7% selected the option "other" when asked about their gender.

Procedure [3] In order to have enough data to evaluate whether clustering situations is useful, it is important that we use a method that generates a diverse sample of situations. To this end, we use a method applied in other research that asks participants to describe a situation in their daily lives (e.g. [57], [136]). This retrospective procedure was shown to encourage participants to report on a wide range of situations. We asked participants to think about two situations which occurred during the past weeks which involved one other person, since our focus is on social situations. We specifically asked for situations involving only one other person, since if needed it is possible to control the effect of the relationship with the other person on the situation. However, the approach would work the same way for situations involving multiple other people. We instructed participants to think of situations where a concrete activity took place, and not situations such as "I saw someone in the street and said hello". A positive example was not given in order to avoid priming the participants towards certain situation types. Participants were asked to describe the situations in 3-4 sentences and to focus on describing

[1]The data can be accessed under: https://doi.org/10.4121/12867041
[2]https://www.prolific.co/
[3]The survey questions can be found in Appendix A

the activity, their relation to the other person, as well as how each person behaved in the situation. Furthermore, we instructed participants to try to think of diverse situations, which involved different people and where different activities took place. To check for consistency, participants had to answer four open questions about the situation they just described: when did the situation take place, what was the main activity, where did the situation occur, and what is the role of the other person.

In the second part of the study, participants were presented with a set of statements to measure psychological characteristics of situations, and they were asked how much each statement applies to each of the situations that they had just described. Examples of statements were "A job needs to be done", "Task-oriented thinking is required" etc. The statements were taken from the S8* scale proposed by Rauthmann and Sherman [136]. This is a validated instrument which can be used to measure the DIAMONDS dimensions of a situation. Each dimension is represented by three statements, for an overall total of 24 statements. Participants could indicate their answers on a scale ranging from 1 (not at all) to 7 (totally).

In the last part, participants were presented with a list of personal values, and they were asked on a slider with values from -10 (fully demote) to 10 (fully promote), how much is each value promoted or demoted in each of their two situation. Participants were presented with 21 personal values, which are based on a version of the Schwartz Value Survey [150] which was used on the European Social Survey [151]. Each of the universal value groups is represented by two values, apart from Universalism which is represented by three. In the original survey, each item of the list describes a feature that a person might exhibit (e.g. "She seeks every chance she can to have fun. It is important to her to do things that give her pleasure."), which correspond to a personal value (e.g. "pleasure"). This was done because the aim of the European Social Survey was to explore personal values that people find important, and for that purpose framing values as features of a person was useful. In this study, we want to know how much a value is promoted or demoted in a certain situation, therefore framing values as qualities of a person would not work. For this reason, we presented participants with the underlying value of each item on the list. The only change that was made to the list was to replace the value "National security" with the value "Health", which is also a value from the Security value group. The reason for this is that we believe it is common for people to commonly encounter situations that can affect their health (e.g. sports, choice of food), but we do not expect them to encounter situations that affect national security.

6.5. Results and Discussion

6.5.1. Variety of Situations

Participants reported situations involving a wide range of other people, including a friend (24%), a family member (20%), a co-worker or supervisor (17%), a romantic partner (12%), an acquaintance (3%) or other (24%, mostly consisting of strangers). These situations comprised a high variety of activities, ranging from work meetings to dinner dates, from sport activities to discussions with other drivers, and everything in between. This is also shown by the high variety of the ratings that participants gave to the psychological characteristics of these situations. The rating for each dimension was calculated as the average score that the participant gave to the three statements representing that dimension, following the guidelines of the S8* measurement scale that we are adopting [136]. As seen in Figure 6.2, most of the dimensions have ratings across the whole range of possible alternatives, apart from Adversity and Mating which tend to have a more confined distribution and less variety in general. The score for each dimension is calculated as the average score across the three statements of the questionnaire that define that dimension. We provide a detailed distribution of answers for each psychological characteristic in Figure 6.2, since this insight will be used to form the subspaces in Section 6.5.2.

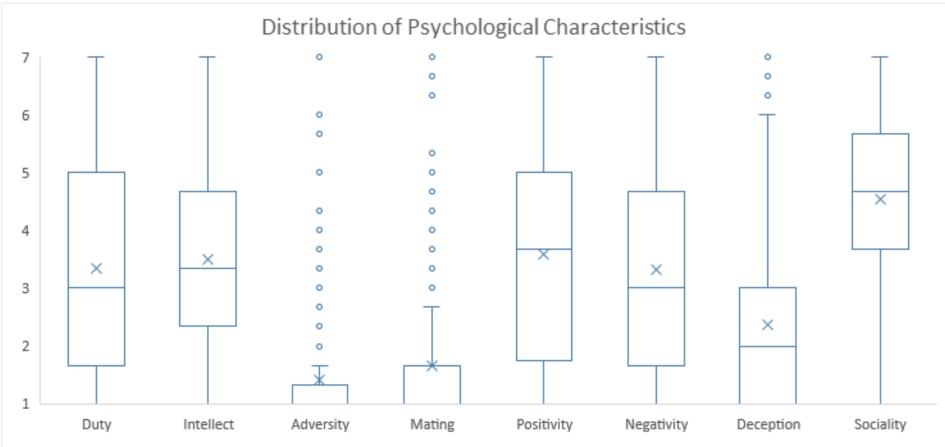

Figure 6.2: Distribution of scores across situations for each dimension, expressing the variety of situations from a point of view of their psychological characteristics. For each boxplot, the middle line represents the median, the sides of the boxes represent the first and third quartiles, and the whiskers represent the minimum and maximum values without considering outliers (which are represented by round points). The x represents the mean scores of the dimensions.

When it comes to personal values that are afforded in these situations according to the participants, the scores also have high variety, as depicted in the distribution presented in Figure 6.3. This distribution suggests that that values were differently promoted or demoted across situations. However, it also holds that most values were slightly promoted on average (overall mean=1.24, SD=4.68). This is in line with research on personal values [143] which views them as positive concepts.

6.5.2. FORMING SITUATION SUBSPACES

In this subsection, we group situations according to their psychological characteristics into situation subspaces. We try an automatic approach, as well as one based on domain knowledge and insights from the data.

AUTOMATIC CLUSTERING

The most straightforward way to form the situation subspaces is by using a clustering algorithm. We tried state of the art algorithms such as K-Means, Affinity Propagation and Agglomerative Clustering using different parameters. The algorithm would receive as an

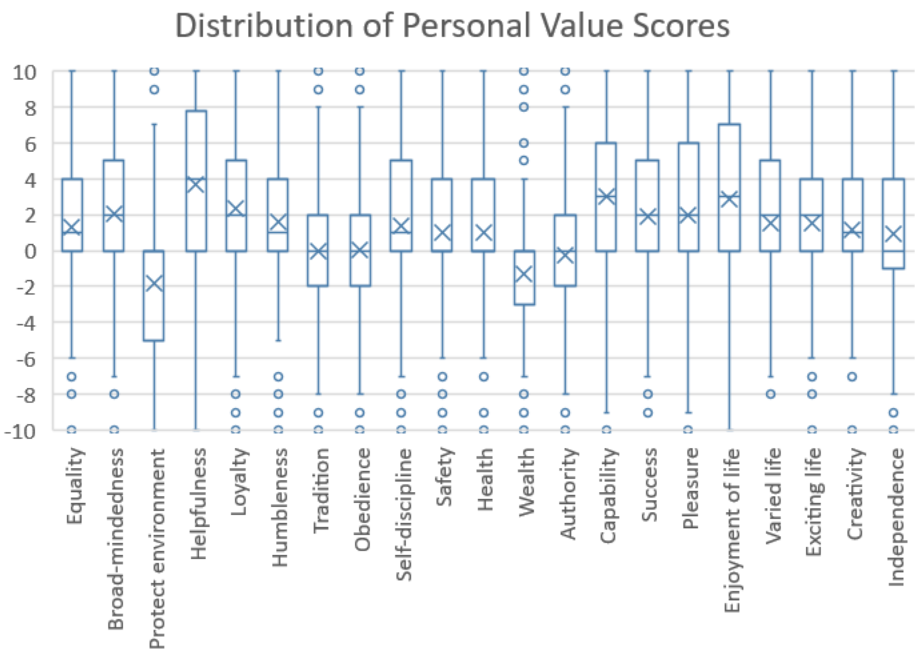

Figure 6.3: Distribution of scores for each personal value across situations.

input the psychological characteristics scores of each situation, and return the cluster to which that situation should belong. We evaluated them with standard metrics used in cases where there is no ground truth when it comes to cluster memberships, such as the Silhouette coefficient and the Davies-Bouldin Index. We used the implementations from the scikit-learn package [128] in Python. The best configuration was achieved by the K-Means algorithm with two clusters, which achieved a Silhouette score of 2.4, and a Davies-Bouldin index of 1.59. These metrics suggest that the data is not well separable when we use all the dimensions in order to perform the clustering. This was to some extent to be expected, considering the high variety of situations, and the fact that there are 8 dimensions and only 300 situations in total. In future work we will collect more situations and explore whether that leads to a higher number of similar situations in the dataset, which could potentially lead to better defined clusters.

While exploring the scores of the dimensions in these two clusters, we notice that in the first cluster Positivity and Mating have a higher score than the average and the other six dimensions have a lower score. In the second cluster this trend is inverted. However, we also notice that each cluster contains situations with scores across the full range of scores for each of the dimensions. First of all, this suggests that these clusters are difficult to interpret/explain since they do not have clear distinguishing features. Secondly, in order to be able to use the Context Space Theory framework, attributes need to have a defined range, which means for at least some of the dimensions we need to have a cutting threshold. This is not the case for the formed clusters, and when faced with a new situation, it is not trivial to determine to which cluster it belongs. Overall, we notice that performing automatic clustering on our data leads to clusters consisting of situations which share some similarity in terms of psychological characteristics, but the division is not granular enough.

Using Data Insight and Domain Knowledge

The next approach will be to use insights from the data as well as domain knowledge in order to manually group situations into situation subspaces. It is important to notice that by "data insights" we only refer to the scores given to the situation dimensions, and not the scores assigned to personal values. From the previous subsection, we learn that trying to cluster over all dimensions is not effective because of the low amount of data and its high variety. For this reason, we use less dimensions in order to define each situation subspace. In order to identify these dimensions, first of all we explore the data. In Figure 6.2 we notice that the dimensions which bring the highest variety to the data are positivity, negativity, intellect and duty. This makes combinations of these dimensions suitable for defining the situation subspaces, since their scores have a high range, and the combinations would lead to subspaces with similar numbers of situations in them. Another insight from the data is that adversity has a very low variety, which makes the situations with a high adversity to form a particular group when compared to the rest. The same applies to mating, but adversity serves the purpose more since it contains outliers. Domain knowledge about the nature of these dimensions can also inform the process of selecting dimensions used to define subspaces. Positivity and negativity, despite being independent concepts, have an inherently opposite flavor. On the other hand, negativity has similar connotations with deception. This is also confirmed by the Pearson correlation coefficients between the data (positivity-negativity: -0.56, negativity-deception: 0.37). This information was used to define six situation subspaces:

- Subspace 1 - High Duty, High Intellect, Low Adversity;
- Subspace 2 - High Positivity, Low Duty, Low Intellect;
- Subspace 3 - High Duty, Low Intellect;
- Subspace 4 - High Adversity;
- Subspace 5 - High Negativity, Low Positivity, Low Duty, Low Intellect, Low Adversity;
- Subspace 6 - High Intellect, Low Duty.

The description "High" refers to scores between 4-7, while the description "Low" refers to scores between 1-3.99 (non-integer scores are possible since each dimension is calculated as the mean of three items from the survey). That means the dimension is highly or lowly characteristic of situations in that subspace. These subspaces allow us to classify 262 out of the 300 situations in our data set. When exploring the remaining situations, we notice that all dimensions other than sociality have a low score. For this reason, we use sociality as a dimension to define the final split, thus forming the last two subspaces:

- Subspace 7 - Low Sociality, and all other dimensions also Low;

- Subspace 8 - High Sociality, and all other dimensions Low.

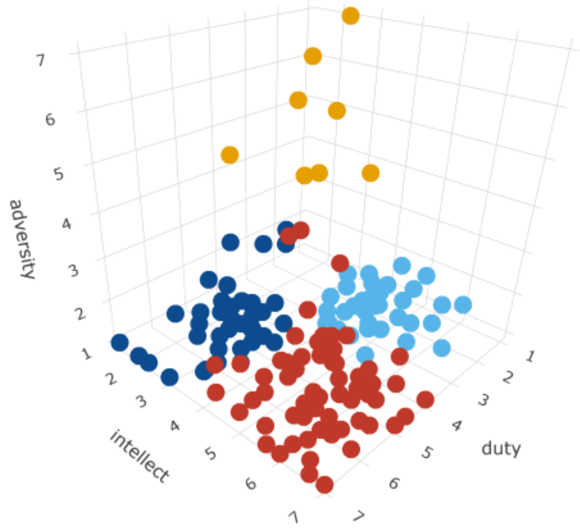

Figure 6.4: Visualisation of four situation subspaces defined by Adversity, Intellect and Duty. Red dots represent situations from Subsp. 1, dark blue dots represent situations from Subsp. 3, orange dots represent situations from Subsp. 4, and light blue dots represent situations from Subsp. 6.

These subspaces are designed to work well with the Context Space Theory framework, since each of them is defined by a set of attributes with specific values. This allows for a straightforward way for classifying a new situation to a subspace. Figure 6.4 provides a visualisation of this, by depicting four of the subspaces projected onto their defining dimensions, for illustration purposes. These defining dimensions enable the subspaces to be more interpretable and explainable in terms of the psychological characteristics that apply to their situations, when compared to the automatic clusters that were created.

We notice that the subspaces are not strictly disjoint. However, this is not a restriction from Context Space Theory, where our approach is based. This also works on an intuitive level, since situations are fluid concepts which can be “in between" two different subspaces. In future work, we will work on strategies on how to break possible ties. Padovitz et al. [123] propose using optional attributes which would increase the probability of a situation being in a subspace.

Using intrinsic metrics for evaluating clusters like we did for the automatic clusters (Silhouette score, Davis-Bouldin Index) would heavily penalize the manual subspaces, since these scores apply to all eight dimensions, whereas the subspaces were defined on a smaller subset of dimensions. For example, in Figure 6.4 we see that the subspaces would be well separated if we only consider the dimensions on which they were defined.

Table 6.1: Distribution of the other person's roles in the situations of each subspace (in percentage). n represents the number of situations (and therefore, the number of people, since situations involve the user and one other person) in each subspace. Fam = Family Member, Rom = Romantic Partner, Fr = Friend, Coll = Colleague, Gr = Group Member

	Fam	Rom	Fr	Coll	Gr	Other
Subspace 1 (n=74)	12.5	9.72	16.67	37.5	4.17	19.44
Subspace 2 (n=77)	23.08	14.1	34.62	8.97	1.28	17.95
Subspace 3 (n=44)	20.45	9.09	20.45	11.36	4.55	34.09
Subspace 4 (n=10)	12.5	0	12.5	25	25	25
Subspace 5 (n=19)	45	15	15	5	5	15
Subspace 6 (n=40)	12.5	12.5	35	15	0	25
Subspace 7 (n=24)	18.52	18.52	11.11	7.41	3.7	40.74
Subspace 8 (n=12)	36.36	9.09	27.27	9.09	0	18.18
All situations (n=300)	20	12	24	17	3.33	23.67

In future work it will be important to define evaluation metrics for manually created subspaces.

We notice a high diversity of activities taking place in the situations of each subspace. For example, Subspace 1 (defined by high duty, high intellect and low adversity), comprises, apart from work situations, also activities such as going to a suture course with a friend, or discussing the family finances with the partner. Similarly, Subspace 4 (defined by high adversity) includes situations ranging from someone being accused of cheating in a card game, to someone being lectured from the CEO of the company. This supports our initial premise that analysing the psychological characteristics of situations can point out to similarities between situations that seem very different at first sight. A similar variety is also present when it comes to the role of the other person in the situation. In our setup, roles are mutually exclusive. The distributions are depicted in Table 6.1. As we can see, in each subspace there are people from almost all the roles present. As expected, Subspace 1 (situations with high intellect and duty) include more colleagues, and Subspace 2 (situations with high positivity, low duty and low intellect) include more family and friends, and less colleagues. This aspect will be analysed further in future work.

6.5.3. Promoted and Demoted Personal Values

In this section, we explore whether specific values tend to be more promoted or demoted across situation subspaces. We look at this from two points of view. First of all, we take into consideration statistical significance. For this, we perform the Wilcoxon rank-sum test to check whether the scores of each value in the situations of a subspace are significantly different from the ones in the rest of situations. Secondly, we look at the mean scores. We consider that a subspace strongly promotes a value when the mean score of the values in its situations is higher than 3.5, and it strongly demotes a value when the mean score is lower than -2.5. Demoting has a lower threshold since we notice that participants tend to give slightly more positive scores overall (the overall mean is 1.24).

6

Table 6.2: Average score for each value in each cluster as well as the full data set. Scores in bold mean that the value is promoted or demoted in that cluster, with boundaries at <-2.5 for demoting and >3.5 for promoting. Scores marked with * suggest statistical significance with p<0.05 when performing the unpaired Wilcoxon rank-sum test for the cluster vs. the rest of the data.

Value (value group)	Subsp 1	Subsp 2	Subsp 3	Subsp 4	Subsp 5	Subsp 6	Subsp 7	Subsp 8	All Sit.
Equality (Universality)	2.2	1.72	1.11	-1.5*	**-2.63***	2.03	0.96	1.82	1.32
Broad-mindedness (Universality)	**3.5***	1.74	1.07*	-0.5*	1	**3.98***	-0.37*	1.36	2.07
Protect environment (Universality)	-2.04	-2.7	-0.95	-1.88	-2.37	-1.25	-0.52	-1.36	-1.79
Helpfulness (Benevolence)	**5.58***	2.5*	**4.41**	-1.5*	0.63*	3.48	2.89	**6.18**	**3.66**
Loyalty (Benevolence)	3.07	3.26	1.8	-2.38*	0.37*	2.78	0.33*	3.45	2.33
Humbleness (Tradition)	2.47	2.05	1.34	-0.75*	-0.63	1.68	1.07	1.09	1.64
Tradition (Tradition)	0.45	-0.09	0.25	-0.88	**-3.05***	0.85	-0.93	1.45	-0.04
Obedience (Conformity)	1.49*	-0.79	0.52	**-2.63**	2	-1.15	-1.11	0.55	0.05
Self-discipline (Conformity)	**3.68***	-1.18*	2.82*	1.5	1.68	1.18	1.33	1	1.39
Safety (Security)	1.95	-0.21*	2.3	**-3.88***	0.11	1.4	0.78	2.36	1.02
Health (Security)	1.18	0.2*	1.89	-1	-0.32	1.8	1.33	**3.91**	1.01
Wealth (Power)	-0.89	-1.55	-1.32	-0.88	-1.26	-1.48	-1.63	-0.09	-1.28
Authority (Power)	1.27*	-1.86*	1.34*	-1	-1.47	-0.48	-1.3	1	-0.24
Capability (Achievement)	**5.45***	1.78*	**3.86**	1	0.74*	3.15	1.11*	2.09	2.99
Success (Achievement)	**4.04***	1.29	2.55	0.63	-1.63*	1.83	1.19	0.82	1.93
Pleasure (Hedonism)	1.15	**5.76***	-0.77*	**-3.5***	**-3.63***	**4.55***	0.3	0.18	1.94
Enjoyment of life (Hedonism)	1.93	**6.82***	0.02*	**-3.25***	-0.63*	**4.73***	1.15*	2.45	2.9
A varied life (Stimulation)	1.7	2.62*	1.5	-0.63	-0.05	2.33	-1.04*	1.82	1.56
An exciting life (Stimulation)	0.85	**4.01***	0*	-1.38*	-0.05	2.58	-0.26*	-0.18*	1.5
Creativity (Self-Direction)	2.68*	1.54	0.39	-1.13	**-2.74***	2.15	-0.96	1.64	1.18
Independence (Self-Direction)	2.39*	0.33	1.39	-0.88	-1.53	0.65	0.37	1.64	0.91

Despite the distributions not being strictly normal, we believe the mean can be informative since the scale is limited between -10 and 10 so there are no values that can greatly skew it. We also calculated the median, and there is a very high overlap in the values that fulfill the criteria (22 out of 26). We do not report the medians for space purposes. We perform this analysis for the automatically created clusters, as well as for our manually formed subspaces.

When it comes to the automatically created clusters, we notice that the first one significantly promotes the values pleasure (3.87) and enjoyment of life (4.87), whereas the second cluster significantly promotes the value capability (4.08). No values are significantly demoted in either cluster. We do not report all values for space purposes. When comparing these results to the interpretation of the clusters using the psychological characteristics of situations, it seems intuitive that the cluster with higher positivity and mating promotes pleasure and enjoyment of life, whereas the cluster with higher duty and intellect promotes capability. The divisions are not granular enough to help us determine a larger number of promoted and demoted values, since we have only two clusters which consist of diverse situations. However, this analysis hints towards the idea that subsets of the data which share similar psychological characteristics do tend to promote certain values more than others, when compared to the overall data.

Next, we perform the same analysis for our manually crafted situation subspaces (Table 6.2). We notice that 5 of the subspaces significantly promote or demote some personal values, thus supporting our initial hypothesis. By analysing these results further, we notice that they are also aligned with the common sense understanding of these concepts: values such as pleasure and enjoyment of life are promoted in situations defined by high positivity (Subspace 2) and demoted in situations defined by high adversity (Subspace 4). Moreover, situations defined by high intellect and duty promote values such as helpfulness, capability and success. These intuitive connections suggest that a support agent that uses this method would have the possibility to explain its suggestions to the users in an understandable way. Furthermore, it seems like the promoted or demoted values are affected by the combination of dimensions, rather than by each dimension individually. For instance, situations defined by both high intellect and duty (Subspace 1) significantly promote success and helpfulness, whereas situations defined by high duty and low intellect (Subspace 3) or low duty and high intellect (Subspace 6) do not promote these values.

6.6. CONCLUSION

6.6.1. CONTRIBUTIONS

In this chapter we present an approach in which we group situations into subspaces by using their psychological characteristics as attributes, and show that these subspaces can be used to determine which personal values are promoted or demoted in these situations. In order to explore our research question, we use automatic clustering, as well as insights from the data combined with domain knowledge, in order to group situations according to their psychological characteristics. We notice that automatic methods lead to clusters which are not well defined, while the manual method allowed us to form groups that fit the requirements of Context Space Theory.

Secondly, we show that certain personal values are significantly more promoted or demoted in specific situation subspaces, thus confirming our research hypothesis. This can be used as a method to automatically determine how the situation that a user faces affects the personal values of the user. This would be a useful extension for current support agents [31], [161] that rely only on information from the users to know the effect it has on personal values.

An advantage of this approach is its potential for providing explainable support to the user. Our methods are inherently more explainable than black box approaches, and we borrow the attributes that form the basis of our approach from social psychology. Concepts such as the psychological characteristics or personal values are potentially more understandable for users.

6.6.2. LIMITATIONS AND FUTURE WORK

Considering that the work is still in its early stage, there are limitations which we aim to tackle in the future. First of all, we assume that we already know the psychological characteristics of a situation. This is not a trivial task, and in order to have a supportive agent that can help in real life cases, these characteristics will have to be inferred from situation cues. Chapter 3 provides initial evidence that they can be used to infer concepts such as the priority of situations. In the future, we will explore whether that approach can be applied to the psychological characteristics of situations.

Secondly, we detect more affected values in the manually defined situation subspaces. While this approach is not necessarily weaker than an automatic approach, it has to be tested with a wider range of situations. The reason for this is that it was crafted particularly for this set of situations, so its effectiveness for another set of situations is to be determined. In the future, we will work on having a well-defined formal procedure on how to form situation subspaces by using the psychological characteristics of situations as context attributes. Another option will be to explore forming automatic clusters by considering a subset of the dimensions.

Next, the promoted and demoted values need to be analysed further. We notice three of the subspaces do not promote or demote any personal values, and some personal values are neither promoted nor demoted in any subspace. In future work, we will explore using a more specific list of values which are salient to daily life situations. Lastly, we will explore whether situation subspaces can help determine concepts other than personal values, such as expected behaviour.

7

Discussion and Conclusion

The research presented in this thesis focuses on enabling support agents to exhibit social situation awareness. As application domain the thesis focuses on agenda management support agents that take into account the social aspect of meetings when making suggestions to users regarding which meeting to attend when different meetings overlap. The thesis answers this overarching research question:

> *What concepts and information processing techniques would enable support agents to exhibit social situation awareness?*

In this section, we start by providing a discussion of how the research questions were tackled, and the results that were obtained. Secondly, we discuss the limitations of our approach. Next, we formulate the scientific and practical contributions of our work, as well as review possible ethical considerations. Lastly, we introduce possible directions for future work based on our findings.

7.1. Discussion

Five sub-questions were introduced in order to capture the different facets of the main research question. The studies presented in this thesis aim to answer these research questions. In this section we discuss the findings from these research questions:

RQ1: What components should a support agent include in order to manifest social situation awareness and how can these components be organized in a conceptual architecture?
This question is tackled in Chapter 2. We started by providing a definition for social situation awareness in support agents. Based on the literature, we identified five requirements that support agents need to fulfill in order to be aware of the social situation of the user. To tackle these requirements we put forward a conceptual architecture inspired by Endsley's model of situation awareness [47], initiated with concepts from social psychology [135]. The proposed architecture (Figure 2.1) consists of three levels: 1)

social situation perception, in which the social situation is modelled from the user's perspective, 2) *social situation comprehension*, in which the social situation is described in terms of its psychological characteristics, and 3) *social situation projection* in which the expected behavior of the user and the promoted and demoted values of the situation are determined. Implementing the levels of this conceptual architecture and transitioning between levels is the focus of the next chapters.

RQ2: Which elements of the user's environment should be modelled in order to represent a present or future social situation of a user?
Chapter 3 focuses on modelling social situations from the point of view of the users (Level 1: social situation perception). To identify the relevant elements of the social environment we explored literature from social sciences focusing on social relationships. Based on the findings, we proposed a two-level ontology for representing social situations. The top level introduces the concepts that need to be represented in order to capture social situations, whereas the bottom level contains the specific features that can be used. Based on the ontology, we introduced a set of social situation features that can be used to model meetings between two people (the user and the other person) from the point of view of the user. These features consist of social background features that describe the social relationship between the user and the other person, and social situation cues that describe other elements of the situation. We conducted a pilot study with twenty subjects, and asked them to describe their relationships with five people from their social circle using the social background model. Subjects found the features highly understandable and moderately representative of their relationships with someone.

RQ3: To what extent can machine learning techniques predict the priority of the situation for the user on the basis of social situation features as input?
The third research question builds upon the findings of the previous one, and is explored in Chapter 4. Our aim was to explore the predictive capabilities of the social situation features when used in agenda management support agents. Specifically, we used social situation features as input in machine learning models to predict the priority of social situations. First, we conducted a pre-study and an extensive literature search in order to update and consolidate the set of social situation features proposed in Chapter 3. Then, we conducted a crowd-sourcing user study where 278 subjects provided information about their relationship with five people from their social circle. Then, the participants were presented with eight hypothetical social situations involving them and one of the people from their social circle, and were asked about the priority level that they would assign to the situation. We used the data to train and test different learning algorithms. The best performing model is a random forest regressor, with a mean absolute error of 1.35 (on a 7 points scale). We noticed that the error was partially caused by the imbalanced data set, and that the model performance was worse for situations with a low level of priority, which is something that needs to be taken into account. Lastly, we showed how we can determine the most informative features that contributed to the predictions, which will serve as a basis for explanations given to users.

RQ4: To what extent can machine learning techniques predict the psychological characteristics of a social situation on the basis of social situation features as input?

In the previous chapter our predictive model went directly from Level 1 to Level 3 of the social situation awareness architecture, and did not explicitly tackle Level 2 (social situation comprehension). This is addressed in Chapter 5. We proposed to use the psychological characteristics of situations as a basis for social situation comprehension in an agenda management support agent. More specifically, we used the DIAMONDS taxonomy proposed by Rauthmann et al. [135]. The first part of the chapter explores the usefulness of psychological characteristics of situations as predictors of the priority of social situation. In this chapter, we used the data collected in Chapter 4 enriched with information about the psychological characteristics of each situation from the point of view of the users. First, we showed that using psychological characteristics of situations as input to a random forest regressor leads to a better prediction of the priority of social situations than when using as input social situation features. Secondly, we showed that it is possible to use machine learning models to automatically predict the psychological characteristics of a situation by using as input social situation features. In order to evaluate the role of social situation features (Level 1 information) and psychological characteristics of situations (Level 2 information) beyond their predictive power, we assessed their role as a basis for explanations given by support agents. We conducted two user studies with a total of 290 subjects, who were presented with explanations given by a hypothetical agenda management agent regarding reasons why a meeting was selected when two meetings were overlapping. The studies show that explanations based on Level 1 and Level 2 information were complete, satisfying, and likely to convince users. Furthermore, results show that whether subjects prefer explanations based on Level 1 or Level 2 depended on the type of situation, thus showing the need for both.

RQ5: How can the promoted or demoted values of a social situation be determined on the basis of the psychological characteristics of that situation?

Chapter 6 tackles our last research question, and deals with social situation projection. Specifically, it focuses on assessing which personal values are promoted or demoted in social situations. For this chapter, we conducted a user study in which we asked 150 participants to describe two social situations involving one other person who they interacted with in the past week. Then, for each situation subjects answered a survey about its psychological characteristics (based on the DIAMONDS taxonomy [135]), as well as a survey about whether the situation promotes, demotes or does not affect specific personal values. As a first step, we grouped similar situations into situation subspaces based on their psychological characteristics using insight from the data and domain knowledge. Results showed that specific personal values are more promoted or demoted in some situation subspaces. For instance, situations with high duty and high intellect promote values helpfulness, self-discipline, capability and success, while situations with a high level of adversity demote the values safety, pleasure and enjoyment of life. These insights serve as a basis for automatically assessing the promoted and demoted values of a social situation based on the psychological characteristics of the situation.

7.2. LIMITATIONS

To fully appreciate the findings presented in this thesis, it is important to consider the limitations of the described studies. The first limitation is related to our application domain. Our instantiations were based on a socially aware agenda management support agent, and the modelled concepts were related to this domain. This makes it difficult to generalize our findings to other types of support agents. This is particularly relevant to social situation projection, since the predicted behaviors (in our case, the priority level of meetings) are directly related to the domain. However, this limitation only concerns our evaluation, and not necessarily the applicability of the approach as a whole. We expect that the proposed social situation awareness architecture can be adapted for different application domains. For instance, stress management support has been identified as a domain in which information about the social situation would be crucial [102]. Our proposed Level 1 and Level 2 concepts would be suitable for that domain, particularly features such as depth of acquaintance and quality of relationship and psychological characteristics such as adversity and negativity. Level 3 information needs to be adapted based on existing work on stress and its behavioral manifestation. Furthermore, user studies need to be conducted in order to collect data to build predictive models needed to transition between the different levels.

Secondly, the studies presented in this thesis were conducted online through crowd sourcing platforms. Due to this, studies had to be shorter and more complex situations had to be avoided. Furthermore, in some of the studies subjects were presented with hypothetical scenarios. This choice was made in order to ensure a wider variety of situations, however there are no guarantees whether real-life situations would generate the same results. Another limitation brought by the online nature of the studies was the fact that we collected data regarding a small number situations for each participant, and the predictive models were built on aggregated data. While this is not necessarily a limitation, it is to be tested how the models would perform when tested on individual users.

Thirdly, in this thesis we focused on the levels of the social situation awareness architecture individually, and transitioning between levels. This was a natural first step, since it allowed us to draw conclusions regarding the feasibility of the architecture. Our results indicate that it is possible to transition between individual levels, however all the steps were not integrated into one system. Results presented in Chapter 5 showed that the transition between the three levels in the same system is feasible.

Lastly, our user studies focused on modelling dyadic social situations, i.e. situations involving two people: the user and another person. This was done to simplify the process of modelling social relationships. Covering social situations involving more people would not require changes to the components of the conceptual social situation awareness architecture, however we expect changes in the instantiation of these components. This is especially relevant for social situation perception (Level 1), which would need to include further features regarding the social dynamics of the group.

7.3. Conclusion and Contributions

Tackling the research questions proposed in this thesis required developing a conceptual architecture, an ontology, a number of predictive models, collecting data to train the models, as well as measurement tools for several user studies. The next paragraphs highlight the main contributions of the thesis.

7.3.1. SCIENTIFIC CONTRIBUTIONS

In this thesis we contribute to research on support agents by providing a definition of social situation awareness in support agents as well as a set of requirements that support agents have to fulfill in order to be considered social situation aware. These requirements are tackled through the proposed conceptual architecture for social situation awareness in support agents. To the best of our knowledge, this is the first attempt to explicitly take into account social situations in situation awareness. To do so, we combine existing work on situation awareness [47] and social sciences [135]. The resulting conceptual architecture provides a description of the high level components needed to tackle the requirements for social situation awareness in support agents, as well as how these components should be interrelated.

Another contribution of the thesis is the translation into a computational model of different concepts from social science research used to implement the different levels of the proposed architecture. The ontology of social situations introduced in Chapter 3 can serve as an initial point for researchers aiming to model social situations, and it can be extended to account for different types of social situation information. Furthermore, this thesis contributes with an instantiation of the ontology for social relationships of meetings involving two people. This involved extensively reviewing social science literature on social relationships. The derived list of social situation features can be used as a baseline for future research. Another crucial contribution is the use of psychological characteristics of situations for social situation comprehension. Recent research in AI has included insights about the psychological characteristics of the user's personality coming from social science such as the Big Five taxonomy [61] to tackle issues such as user modelling for recommender systems (e.g., [114], [146]). However, to the best of our knowledge, AI research has yet to systematically include the recent insights from social science about the psychological characteristics of situations (e.g., [126], [135]). We contribute to this by showing how one of the existing taxonomies of psychological characteristics of situations, namely the DIAMONDS taxonomy [135], can be used as a basis for social situation comprehension. Lastly, we explored how the concept of personal values can be integrated in support agents, and specifically showed how it can be used for social situation projection.

In order to evaluate the proposed concepts, several user studies were conducted to collect data. Through them, we contribute to the scientific community by providing a set of hypothetical and real social situations as well as different data sets that can be used by researchers to investigate research questions related to social situations, priority of situations, and personal values. Furthermore, different lessons were learned about how to set up such studies. For instance, designers of user studies should instruct users to provide information about diverse types of situations, because not having such guidance can lead to unbalanced data sets.

The predictive models presented throughout the thesis provide the basis for transitioning between the different levels of the architecture. In this thesis, we show how information about a user's social situation can be used to predict the psychological characteristics of that situation, and in turn how the psychological characteristics of a situation can be used to predict the priority of that situation and the personal values that the situation promotes or demotes. The evaluation of these models forms the first benchmark in these novel tasks.

Lastly, in Chapter 5 we show how the concepts introduced in our architecture can be used as a basis for explanations that are satisfying, containing sufficient information, complete, and convincing. This suggests that our proposed three-level architecture and specific social science concepts can contribute to explainable support agents. This is also corroborated by existing work [148] which proposes a situation awareness-based framework for the design of explainable AI.

7.3.2. Practical Contributions

Apart from the scientific community, different groups also benefit from the work presented in this thesis, as shown below:

Developers of Support Agents

Our thesis focuses on improving the support of support agents by allowing them to take into account the social situation of the user. Our proposed architecture in Chapter 2 can serve as a blueprint for designers and developers of support agents. The other chapters present useful insights on the concepts and learning techniques that can be used to implement the proposed concepts. Following our guidelines and fulfilling the requirements proposed in Chapter 2 would allow developers to incorporate information about the social situation of the user in the support mechanism, which could result in a better product.

End Users

Once implemented and integrated in existing personal assistant agents, our proposed agenda management support agent would be beneficial to people with busy schedules who have to spend a lot of time to organize their commitments. The agent could determine the priority of each meeting from the point of view of the user, as well as assess the promoted and demoted personal values, and present the information to the user in conjunction with suggestions regarding which meetings can be postponed or canceled. This would provide more comprehensive support, as opposed to existing systems that base its suggestions mainly on slot availability. An adapted version of our envisioned agent would also be beneficial to coordinators or human personal assistants who have to deal with scheduling meetings for multiple people.

7.4. Ethical Considerations

Before implementing and deploying intelligent systems such as the one described in this thesis, it is crucial to take into account ethical considerations related to their potential impact.

The first ethical consideration is related to the proposed application domain. Having support agents assist people in managing their agenda can be beneficial, but can also have unwanted consequences. For instance, cancelling specific meetings might cause misunderstandings between the user and people from their social circle, or in more extreme cases it might negative repercussions on the user's work or personal life. To mitigate this risk, in our proposed approach control stays in the hands of the user and the agent is limited to only making suggestions to the user rather than making decisions on their behalf.

The second ethical consideration relates to the reflective nature of the approach. The support agent explicitly reasons about the psychological characteristics of the social situation (Level 2), as well as the personal values that the social situation promotes or demotes (Level 3). This information can be conveyed to the user in the form of explanations in order to make the support process transparent. However, this also leads to the user more deliberately and explicitly thinking about their social interactions. While in principle this can be a positive thing, instances can arise in which the user is notified that a social situation involving a person close to the user has a high level of adversity or negativity, or demotes values that are important to the user. This can be caused by wrong inferences in the reasoning process of the agent, or in some cases by the fact that the user is simply not explicitly aware of the nature of those social interactions. Both cases, while not being negative per se, can cause distress to the user, which needs to be taken into account.

Lastly, the use of machine learning models to transition between the levels of the architecture requires consideration. Existing work shows that predictions of machine learning models often involve bias caused by the learning algorithm, the data collection/generation, as well as by existing biases in the world [27], [70]. For this reason, a more in-depth analysis needs to be conducted before these aspects can be deployed. In our proposed architecture data is elicited directly from users rather than collected solely from sensors or other already available data such as social media. This avoids some possible data collection biases, however on the other hand it can lead to reinforcing existing user biases. Furthermore, when building models from aggregated data it would be important to analyze how well predictions work for groups that are underrepresented in the training data. Having some degree of personalization would help predictions be more accurate for individual users. To mitigate some of the risks, we propose the explanation and feedback modules which contribute to more transparency and control.

7.5. Future Work

The work presented in this thesis puts forward a vision for enabling support agents to exhibit social situation awareness. Our results support the feasibility of such a vision in the domain of agenda management support agents. This opens the possibility for new research towards fully social situation aware support agents.

The first step would be to integrate the different levels of the social situation awareness architecture into a prototype system. This would first of all allow to then consolidate and further expand our findings by conducting application-grounded user studies [40] in which participants interact with an agent over a long period of time and provide information for a large number of their daily life social situations. Secondly, it would be important to extend the architecture towards different application domains. This would require further research into how the social situation awareness would have to be instantiated, and what concepts would need to be modelled.

To enable the implementation of application-grounded user studies, more research is needed on elicitation techniques and human-computer interaction. While eliciting information from users on real life social situations is crucial, it is also important to have smooth interactions and to ensure that the system is not overly invasive. This requires combining and further expanding on existing research on information elicitation from users (e.g., [172]) as well as research on using active learning for personalization of support agents while minimizing the needed interaction with users (e.g., [118]).

Another important future direction is related to integrating the social situation awareness component with the support component of the agent. Level 3 of the architecture would provide insight regarding the expected behavior of the user as well as regarding the promoted and demoted personal values of a social situation. This information can inform the support actions of the agents, however more research is needed on the reasoning mechanisms that would allow this, as well as on the practical benefits of such an approach. This requires implementing such a reasoning module in support agents and testing it in user studies, as done in [83].

In this thesis, we showed how the proposed concepts can be used as a basis for explanations given to users. It would be important to combine efforts to improve the explainability of the agent with the implementation of a feedback module which would allow the user to tell the agent when it made a mistake. The explainability step would be the basis for this procedure [71], [98]. Research is needed on ways to effectively include this information to the reasoning mechanisms of the agent.

Lastly, in this thesis when modelling social situations we focused on social relationship features. However, there is also a vast body of research in social signal processing focusing on social interaction features such as body language and their connection to user behavior (e.g., [24], [124], [144]). Integrating this research with our social situation awareness architecture would allow for a more comprehensive representation of social situations. This way the agent would have both information about social relationships elicited from the users, as well as information about the specific social interaction acquired through multimodal sensing. Integrating such concepts in our proposed architecture would be a natural extension, and existing work suggests that information from social interactions can be used to infer psychological characteristics of the situation such as interdependence [41].

References

[1] N. Ajmeri, H. Guo, P. K. Murukannaiah, and M. P. Singh, "Designing Ethical Personal Agents", *IEEE Internet Computing*, vol. 22, no. 2, pp. 16–22, 2018.

[2] N. Ajmeri, P. K. Murukannaiah, H. Guo, and M. P. Singh, "Arnor: Modeling Social Intelligence via Norms to Engineer Privacy-aware Personal Agents", in *Proceedings of the 16th Conference on Autonomous Agents and Multiagent Systems*, International Foundation for Autonomous Agents and Multiagent Systems, 2017, pp. 230–238.

[3] Z. Akata, D. Balliet, M. de Rijke, *et al.*, "A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect with Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence", *Computer*, vol. 53, no. 8, pp. 18–28, 2020.

[4] V. Akman and M. Surav, "Steps Toward Formalizing Context", *AI magazine*, vol. 17, no. 3, pp. 55–72, 1996.

[5] R. Al-Hmouz, W. Pedrycz, and A. Balamash, "Description and Prediction of Time Series: A General Framework of Granular Computing", *Expert Systems with Applications*, vol. 42, no. 10, pp. 4830–4839, 2015.

[6] U. Alegre, J. C. Augusto, and T. Clark, "Engineering Context-aware Systems and Applications: A Survey", *Journal of Systems and Software*, vol. 117, pp. 55–83, 2016.

[7] C. B. Anagnostopoulos, Y. Ntarladimas, and S. Hadjiefthymiades, "Situational Computing: An Innovative Architecture with Imprecise Reasoning", *Journal of Systems and Software*, vol. 80, no. 12, pp. 1993–2014, 2007.

[8] T. C. Antonucci, K. J. Ajrouch, and K. S. Birditt, "The Convoy Model: Explaining Social Relations from a Multidisciplinary Perspective", *The Gerontologist*, vol. 54, no. 1, pp. 82–92, 2013.

[9] T. C. Antonucci and H. Akiyama, "Social Networks in Adult Life and a Preliminary Examination of the Convoy Model", *Journal of Gerontology*, vol. 42, no. 5, pp. 519–527, 1987.

[10] T. Balke, C. da Costa Pereira, F. Dignum, *et al.*, "Norms in MAS: Definitions and Related Concepts", in *Normative Multi-Agent Systems*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, vol. 4, 2013.

[11] J. Barwise, *Situations and Small Worlds*. Standford: CSLI Publications, 1987.

[12] J. Barwise and J. Perry, "Situations and Attitudes", *The Journal of Philosophy*, vol. 78, no. 11, pp. 668–691, 1981.

[13] N. Baumgartner and W. Retschitzegger, "A Survey of Upper Ontologies for Situation Awareness", in *Proceedings of the 4th IASTED International Conference on Knowledge Sharing and Collaborative Engineering*, ACTA Press, 2006, pp. 1–9.

[14] F. Bellifemine, A. Poggi, and G. Rimassa, "JADE: A FIPA-compliant Agent Framework", in *Proceedings of the Practical Applications of Intelligent Agents*, The Practical Application Company Ltd., vol. 99, 1999, pp. 97–108.

[15] G. Benincasa, G. D'Aniello, C. D. Maio, V. Loia, and F. Orciuoli, "Towards Perception-oriented Situation Awareness Systems", in *Proceedings of the 7th IEEE International Conference Intelligent Systems*, Springer, 2015, pp. 813–824.

[16] J. Berka, J. Balata, C. M. Jonker, Z. Mikovec, M. B. van Riemsdijk, and M. L. Tielman, "Misalignment in Semantic User Model Elicitation via Conversational Agents: A Case Study in Navigation Support for Visually Impaired People", *International Journal of Human–Computer Interaction*, pp. 1–17, 2022.

[17] C. Bettini, O. Brdiczka, K. Henricksen, *et al.*, "A Survey of Context Modelling and Reasoning Techniques", *Pervasive and Mobile Computing*, vol. 6, no. 2, pp. 161–180, 2010.

[18] C. Boutilier, "A POMDP Formulation of Preference Elicitation Problems", in *Eighteenth National Conference on Artificial Intelligence*, Edmonton, Alberta, Canada: American Association for Artificial Intelligence, 2002, pp. 239–246.

[19] L. Breiman, "Bagging Predictors", *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[20] ——, "Random Forests", *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[21] L. Breiman, J. Friedman, R. Olshen, and C. Stone, *Classification and Regression Trees*. Routledge, 1994.

[22] N. A. Brown, R. Neel, and R. A. Sherman, "Measuring the Evolutionarily Important Goals of Situations: Situational Affordances for Adaptive Problems", *Evolutionary Psychology*, vol. 13, no. 3, pp. 1–15, 2015.

[23] R. S. Burt, "Network Items and the General Social Survey", *Social Networks*, vol. 6, no. 4, pp. 293–339, 1984.

[24] L. Cabrera-Quiros, A. Demetriou, E. Gedik, L. van der Meij, and H. Hung, "The MatchNMingle Dataset: A Novel Multi-sensor Resource for the Analysis of Social Interactions and Group Dynamics in-the-wild During Free-standing Conversations and Speed Dates", *IEEE Transactions on Affective Computing*, vol. 12, no. 1, pp. 113–130, 2018.

[25] N. Cantor, W. Mischel, and J. C. Schwartz, "A Prototype Analysis of Psychological Situations", *Cognitive Psychology*, vol. 14, no. 1, pp. 45–77, 1982.

[26] G. Castellano, M. G. Cimino, A. M. Fanelli, B. Lazzerini, F. Marcelloni, and M. A. Torsello, "A Multi-agent System for Enabling Collaborative Situation Awareness via Position-based Stigmergy and Neuro-fuzzy Learning", *Neurocomputing*, vol. 135, pp. 86–97, 2014.

[27] J. Chakraborty, S. Majumder, and T. Menzies, "Bias in Machine Learning Software: Why? How? What to Do?", in *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ACM, 2021, pp. 429–440.

[28] Y. Chevaleyre, F. Koriche, J. Lang, J. Mengin, and B. Zanuttini, "Learning Ordinal Preferences on Multiattribute Domains: The Case of CP-nets", in *Preference Learning*, Springer, 2010, pp. 273–296.

[29] M. Chromik and M. Schuessler, "A Taxonomy for Human Subject Evaluation of Black-box Explanations in XAI", in *Proceedings of the Workshop on Explainable Smart Systems for Algorithmic Transparency in Emerging Technologies*, CEUR Workshop Proceedings, 2020.

[30] M. G. Cimino, B. Lazzerini, F. Marcelloni, and A. Ciaramella, "An Adaptive Rule-based Approach for Managing Situation Awareness", *Expert Systems with Applications*, vol. 39, no. 12, pp. 10 796–10 811, 2012.

[31] S. Cranefield, M. Winikoff, V. Dignum, and F. Dignum, "No Pizza for You: Value-based Plan Selection in BDI Agents", in *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligen*, International Joint Conferences on Artificial Intelligence, 2017, pp. 178–184.

[32] D. P. Davison, F. M. Wijnen, V. Charisi, J. van der Meij, D. Reidsma, and V. Evers, "Words of Encouragement: How Praise Delivered by a Social Robot Changes Children's Mindset for Learning", *Journal on multimodal user interfaces*, vol. 15, no. 1, pp. 61–76, 2021.

[33] K. Devlin, *Logic and Information*. Cambridge University Press, 1995.

[34] T. G. Dietterich, "Ensemble Methods in Machine Learning", in *International Workshop on Multiple Classifier Systems*, Springer, 2000, pp. 1–15.

[35] F. Dignum, "Interactions as Social Practices: Towards a Formalization", *arXiv preprint arXiv:1809.08751*, 2018.

[36] F. Dignum, R. Prada, and G. J. Hofstede, "From Autistic to Social Agents", in *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, International Foundation for Autonomous Agents and Multiagent Systems, 2014, pp. 1161–1164.

[37] V. Dignum, *A Model for Organizational Interaction: Based on Agents, Founded in Logic*. SIKS PhD Dissertation Series, 2004.

[38] V. Dignum and F. Dignum, "Contextualized Planning Using Social Practices", in *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*, Springer, 2014, pp. 36–52.

[39] A. Dorri, S. S. Kanhere, and R. Jurdak, "Multi-agent systems: A survey", *Ieee Access*, vol. 6, pp. 28 573–28 593, 2018.

[40] F. Doshi-Velez and B. Kim, "Considerations for Evaluation and Generalization in Interpretable Machine Learning", in *Explainable and Interpretable Models in Computer Vision and Machine Learning*, H. J. Escalante, S. Escalera, I. Guyon, *et al.*, Eds. Springer, 2018, pp. 3–17.

[41] B. Dudzik, S. Columbus, T. M. Hrkalovic, D. Balliet, and H. Hung, "Recognizing Perceived Interdependence in Face-to-Face Negotiations through Multimodal Analysis of Nonverbal Behavior", in *Proceedings of the 2021 International Conference on Multimodal Interaction*, ACM, 2021, pp. 121–130.

[42] B. Dudzik, M.-P. Jansen, F. Burger, *et al.*, "Context in Human Emotion Perception for Automatic Affect Detection: A Survey of Audiovisual Databases", in *Proceedings of the 8th International Conference on Affective Computing & Intelligent Interaction*, IEEE, 2019, pp. 206–212.

[43] M. d'Inverno, M. Luck, P. Noriega, J. A. Rodriguez-Aguilar, and C. Sierra, "Weaving a Fabric of Socially Aware Agents", in *PRIMA: International Conference on Principles and Practice of Multi-Agent Systems*, Springer, 2011, pp. 263–274.

[44] J. A. Edwards and A. Templeton, "The Structure of Perceived Qualities of Situations", *European Journal of Social Psychology*, vol. 35, no. 6, pp. 705–723, 2005.

[45] M. Elgarf, N. Calvo-Barajas, P. Alves-Oliveira, *et al.*, ""and Then What Happens?" Promoting Children's Verbal Creativity Using a Robot", in *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, ACM, 2022, pp. 71–79.

[46] M. R. Endsley, "Situation Awareness Global Assessment Technique (SAGAT)", in *Proceedings of the IEEE National Aerospace and Electronics Conference*, IEEE, 1988, pp. 789–795.

[47] ——, "Toward a Theory of Situation Awareness in Dynamic Systems", *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.

[48] M. R. Endsley, "Theoretical Underpinnings of Situation Awareness: A Critical Review", in *Situation Awareness Analysis and Measurement*, M. R. Endsley and D. J. Garland, Eds. CRC Press, 2000, pp. 3–26.

[49] Y.-H. Feng, T.-H. Teng, and A.-H. Tan, "Modelling Situation Awareness for Context-aware Decision Support", *Expert Systems with Applications*, vol. 36, no. 1, 2009.

[50] A. P. Fiske, "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations", *Psychological Review*, vol. 99, no. 4, p. 689, 1992.

[51] K. K. Fitzpatrick, A. Darcy, and M. Vierhile, "Delivering Cognitive Behavior Therapy to Young Adults with Symptoms of Depression and Anxiety Asing a Fully Automated Conversational Agent (Woebot): A Randomized Controlled Trial", *JMIR Mental Health*, vol. 4, no. 2, e19, 2017.

[52] N. Fornara, F. Viganò, and M. Colombetti, "Agent Communication and Artificial Institutions", *Autonomous Agents and Multi-Agent Systems*, vol. 14, no. 2, pp. 121–142, 2007.

[53] B. Friedman, P. H. Kahn, A. Borning, and A. Huldtgren, "Value Sensitive Design and Information Systems", *Early Engagement and New Technologies: Opening U the Laboratory*, vol. 16, pp. 55–95, 2013.

[54] J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine", *Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.

[55] K. K. Fullam, T. Klos, G. Muller, J. Sabater-Mir, K. S. Barber, and L. Vercouter, "The Agent Reputation and Trust (art) Testbed", in *International Conference on Trust Management*, Springer, 2006, pp. 439–442.

[56] J. D. Gehrke, “Evaluating Situation Awareness of Autonomous Systems”, in *Performance evaluation and benchmarking of intelligent systems*, Springer, 2009, pp. 93–111.

[57] F. H. Gerpott, D. Balliet, S. Columbus, C. Molho, and R. E. de Vries, “How Do People Think about Interdependence? A Multidimensional Model of Subjective Outcome Interdependence”, *Journal of Personality and Social Psychology*, vol. 115, no. 4, pp. 716–742, 2018.

[58] A. Ghandeharioun, D. McDuff, M. Czerwinski, and K. Rowan, “Emma: An Emotion-aware Wellbeing Chatbot”, in *Proceedings of the 8th International Conference on Affective Computing and Intelligent Interaction*, IEEE, 2019, pp. 1–7.

[59] G. Gigerenzer, “The Modularity of Social Intelligence”, *Machiavellian intelligence II: Extensions and evaluations*, vol. 2, no. 264, pp. 264–288, 1997.

[60] U. Gnewuch, S. Morana, and A. Maedche, “Towards Designing Cooperative and Social Conversational Agents for Customer Service”, in *Proceedings of the International Conference on Information Systems*, Association for Information Systems, 2017.

[61] L. R. Goldberg, “An Alternative “Description of Personality": The Big-five Factor Structure”, *Journal of Personality and Social Psychology*, vol. 59, no. 6, pp. 1216–1229, 1990.

[62] B. Goodman and S. Flaxman, “European Union Regulations on Algorithmic Decision-making and a “Right to Explanation””, *AI magazine*, vol. 38, no. 3, pp. 50–57, 2017.

[63] M. Granovetter, “The Strength of Weak Ties: A Network Theory Revisited”, *Sociological Theory*, vol. 1, pp. 201–233, 1983.

[64] M. Granovetter, “Weak Ties and Strong Ties”, *American Journal of Sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.

[65] B. Grosz, “What Question Would Turing Pose Today?”, *AI magazine*, vol. 33, no. 4, pp. 73–73, 2012.

[66] S. Gu, B. Kelly, and D. Xiu, “Empirical Asset Pricing via Machine Learning”, National Bureau of Economic Research, Tech. Rep., 2018.

[67] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, “A Survey of Methods for Explaining Black Box Models”, *ACM computing surveys (CSUR)*, vol. 51, no. 5, pp. 1–42, 2018.

[68] D. Gunning and D. Aha, “DARPA’s Explainable Artificial Intelligence (XAI) Program”, *AI magazine*, vol. 40, no. 2, pp. 44–58, 2019.

[69] C. A. Heaney and B. A. Israel, “Social Networks and Social Support”, *Health Behavior and Health Education: Theory, Research, and Practice*, vol. 4, pp. 189–210, 2008.

[70] T. Hellström, V. Dignum, and S. Bensch, “Bias in Mmachine Learning - What Is It Good For?”, in *Proceedings of the 1st International Workshop on New Foundations for Human-Centered AI*, 2020, pp. 3–10.

[71] K. V. Hindriks, “Debugging Is Explaining”, in *International Conference on Principles and Practice of Multi-Agent Systems*, Springer, 2012, pp. 31–45.

[72] R. Hoekstra, “Ontology Representation: Design Patterns and Ontologies that Make Sense”, in *Proceedings of the 2009 Conference on Ontology Representation: Design Patterns and Ontologies that Make Sense*, IOS Press, 2009, pp. 1–236.

[73] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman, “Metrics for Explainable AI: Challenges and Prospects”, *arXiv preprint arXiv:1812.04608*, 2018.

[74] K. T. Horstmann, J. F. Rauthmann, and R. A. Sherman, “The Measurement of Situational Influences”, *The SAGE Handbook of Personality and Individual Differences*, vol. 1, pp. 465–484, 2017.

[75] H.-F. Hsieh and S. E. Shannon, “Three Approaches to Qualitative Content Analysis”, *Qualitative health research*, vol. 15, no. 9, pp. 1277–1288, 2005.

[76] A. Jameson, B. Berendt, S. Gabrielli, *et al.*, “Choice Architecture for Hhuman-computer Interaction”, *Foundations and Trends in Human-Computer Interaction*, vol. 7, no. 1–2, pp. 1–235, 2014.

[77] Y. Javed, T. Norris, and D. Johnston, “Ontology-based Inference to Enhance Team Situation Awareness in Emergency Management”, in *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management*, ISCRAM, 2011.

[78] M. Johnson, J. M. Bradshaw, P. J. Feltovich, C. M. Jonker, M. B. Van Riemsdijk, and M. Sierhuis, “Coactive Design: Designing Support for Interdependence in Joint Activity”, *Journal of Human-Robot Interaction*, vol. 3, no. 1, pp. 43–69, 2014.

[79] R. E. Jones, E. S. Connors, and M. R. Endsley, “A Framework for Representing Agent and Human Situation Awareness”, in *Proceedings of the IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support*, IEEE, 2011, pp. 226–233.

[80] R. L. Kahn and T. C. Antonucci, “Convoys Over the Life Course: Attachment, Roles, and Social Support”, in *Life-span Development and Behavior*, P. B. Baltes and O. G. Brim, Eds. Academic Press, 1980.

[81] G. A. Kaminka, “Curing Robot Autism: A Challenge”, in *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, International Foundation for Autonomous Agents and Multiagent Systems, 2013, pp. 801–804.

[82] A. Kayal, W.-P. Brinkman, R. Gouman, M. A. Neerincx, and M. B. Van Riemsdijk, “A Value-centric Model to Ground Norms and Requirements for e-Partners of Children”, in *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*, Springer, 2013, pp. 329–345.

[83] A. Kayal, W.-P. Brinkman, M. A. Neerincx, and M. B. V. Riemsdijk, “Automatic Resolution of Normative Conflicts in Supportive Technology Based on User Values”, *ACM Transactions on Internet Technology*, vol. 18, no. 4, pp. 1–21, 2018.

[84] V. Kepuska and G. Bohouta, "Next-generation of Virtual Personal Assistants (Microsoft Cortana, Apple Siri, Amazon Alexa and Google Home)", in *IEEE 8th Annual Computing and Communication Workshop and Conference*, IEEE, 2018, pp. 99–103.

[85] M. S. Kließ, C. M. Jonker, and M. B. van Riemsdijk, "A Temporal Logic for Modelling Activities of Daily Living", in *Proceedings of the 25th International Symposium on Temporal Representation and Reasoning*, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018, pp. 1–15.

[86] M. S. Kließ, M. Stoelinga, and M Riemsdijk, "From Good Intentions to Behaviour Change", in *Proceedings of the International Conference on Principles and Practice of Multi-Agent Systems*, Springer, 2019, pp. 354–369.

[87] F. Klügl, "Using the Affordance Concept for Model Design in Agent-based Simulation", *Annals of Mathematics and Artificial Intelligence*, vol. 78, no. 1, pp. 21–44, 2016.

[88] M. M. Kokar, C. J. Matheus, and K. Baclawski, "Ontology-based Situation Awareness", *Information Fusion*, vol. 10, no. 1, pp. 83–98, 2009.

[89] I. Kola, C. M. Jonker, and M. B. van Riemsdijk, "Modelling the Social Environment: Towards Socially Adaptive Electronic Partners", in *Proceedings of the 10th International Workshop on Modelling and Reasoning in Context*, CEUR Workshop Proceedings, 2018, pp. 30–34.

[90] ——, *Pilot Experiment Exploring the Priority of Social Situations*, 2019. DOI: 10.4121/uuid:e18fb318-c1d4-4ccc-9b4f-be48e1ee49e2.

[91] ——, "What Does It Take to Create Social Awareness for Support Agents?", in *Pre-proceedings of the International Workshop on Engineering Multi-Agent Systems*, Springer, 2019.

[92] I. Kola, C. M. Jonker, and M. B. van Riemsdijk, "Who's That? Social Situation Awareness for Behaviour Support Agents", in *International Workshop on Engineering Multi-Agent Systems*, Springer, 2019, pp. 127–151.

[93] I. Kola, C. M. Jonker, and M. B. van Riemsdijk, "Using Psychological Characteristics of Situations for Social Situation Comprehension in Support Agents", *arXiv preprint arXiv:2110.09397*, 2021.

[94] I. Kola, C. M. Jonker, M. L. Tielman, and M. B. van Riemsdijk, "Grouping Situations Based on their Psychological Characteristics Gives Insight into Personal Values", in *Proceedings of the 11th International Workshop on Modelling and Reasoning in Context*, CEUR Workshop Proceedings, 2020, pp. 17–26.

[95] I. Kola, P. K. Murukannaiah, C. M. Jonker, and M. B. Van Riemsdijk, "A Conceptual Architecture for Social Situation Awareness in Support Agents", *IEEE Intelligent Systems*, 2022.

[96] I. Kola, M. L. Tielman, C. M. Jonker, and M. B. van Riemsdijk, "Predicting the Priority of Social Situations for Personal Assistant Agents", in *PRIMA: Principles and Practice of Multi-Agent Systems*, Springer, 2020, pp. 231–247.

[97] R. Kop, M. Hoogendoorn, and M. C. Klein, "A Personalized Support Agent for Depressed Patients: Forecasting Patient Behavior Using a Mood and Coping Model", in *2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies*, IEEE, vol. 3, 2014, pp. 302–309.

[98] T. Kulesza, M. Burnett, W.-K. Wong, and S. Stumpf, "Principles of Explanatory Debugging to Personalize Interactive Machine Learning", in *Proceedings of the 20th International Conference on Intelligent User Interfaces*, ACM, 2015, pp. 126–137.

[99] J. B. Larsen, V. Dignum, J. Villadsen, and F. Dignum, "Querying Social Practices in Hospital Context", in *10th International Conference on Agents and Artificial Intelligence*, SCITEPRESS Digital Library, 2018, pp. 405–412.

[100] C. A. Le Dantec, E. S. Poole, and S. P. Wyche, "Values as Lived Experience: Evolving Value Sensitive Design in Support of Value Discovery", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 1141–1150.

[101] K. Lewin, "Field Theory and Experiment in Social Psychology: Concepts and Methods", *American Journal of Sociology*, vol. 44, no. 6, pp. 868–896, 1939.

[102] X. S. Li, M. C. Rozendaal, K. Jansen, C. Jonker, and E. Vermetten, "Things That Help Out: Designing Smart Wearables as Partners in Stress Management", *AI & SOCIETY*, vol. 36, no. 1, pp. 251–261, 2021.

[103] B. Y. Lim, A. K. Dey, and D. Avrahami, "Why and Why Not Explanations Improve the Intelligibility of Context-aware Intelligent Systems", in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2009, pp. 2119–2128.

[104] V. Loia, G. D'Aniello, A. Gaeta, and F. Orciuoli, "Enforcing Situation Awareness with Granular Computing: A Systematic Overview and New Perspectives", *Granular Computing*, vol. 1, no. 2, pp. 127–143, 2016.

[105] S. W. Loke, "Representing and Reasoning with Situations for Context-aware Pervasive Computing: A Logic Programming Perspective", *The Knowledge Engineering Review*, vol. 19, no. 3, pp. 213–233, 2004.

[106] S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions", in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2017, pp. 4765–4774.

[107] J. A. Maestro-Prieto, S. Rodríguez, R. Casado, and J. M. Corchado, "Agent Organisations: From Independent Agents to Virtual Organisations and Oocieties of Agents", *Advances in Distributed Computing and Artificial Intelligence Journal*, vol. 9, no. 4, pp. 55–70, 2020.

[108] D. Magnusson, "Interactional Psychology: Present Status and Future Prospects", in *Personality at the Cross-roads: Current Issues in International Psychology*. Lawrence Erlbaum Associates, 1977, pp. 3–31.

[109] T. Mainela, "Types and Functions of Social Relationships in the Organizing of an International Joint Venture", *Industrial Marketing Management*, vol. 36, no. 1, pp. 87–98, 2007.

[110] S. Manolios, A. Hanjalic, and C. C. Liem, “The Influence of Personal Values on Music Taste: Towards Value-based Music Recommendations”, in *Proceedings of the 13th ACM Conference on Recommender Systems*, ACM, 2019, pp. 501–505.

[111] O. McAree, J. M. Aitken, and S. M. Veres, “Towards Artificial Situation Awareness by Autonomous Vehicles”, *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 7038–7043, 2017.

[112] E. Meijer, W. A. Gebhardt, A. Dijkstra, M. C. Willemsen, and C. Van Laar, “Quitting Smoking: The Importance of Non-smoker Identity in Predicting Smoking Behaviour and Responses to a Smoking Ban”, *Psychology & Health*, vol. 30, no. 12, pp. 1387–1409, 2015.

[113] E. Meijer, W. A. Gebhardt, C. Van Laar, R. Kawous, and S. C. Beijk, “Socio-economic Status in Relation to Smoking: The Role of (Expected and Desired) Social Support and Quitter Identity”, *Social Science & Medicine*, vol. 162, pp. 41–49, 2016.

[114] A. B. Melchiorre and M. Schedl, “Personality Correlates of Music Audio Preferences for Modelling Music Listeners”, in *Proceedings of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, ACM, 2020, pp. 313–317.

[115] T. Miller, “Explanation in Artificial Intelligence: Insights from the Social Sciences”, *Artificial Intelligence*, vol. 267, pp. 1–38, 2019.

[116] S. T. Mueller, R. R. Hoffman, W. Clancey, A. Emrey, and G. Klein, “Explanation in Human-AI Systems: A Literature Meta-review, Synopsis of Key Ideas and Publications, and Bibliography for Explainable AI”, *arXiv preprint arXiv:1902.01876*, 2019.

[117] P. K. Murukannaiah and M. P. Singh, “Platys Social: Relating Shared Places and Private Social Circles”, *IEEE Internet Computing*, vol. 16, no. 3, pp. 53–59, 2012.

[118] ——, “Platys: An Active Learning Framework for Place-aware Application Development and its Evaluation”, *ACM Transactions on Software Engineering and Methodology*, vol. 24, no. 3, pp. 1–32, 2015.

[119] K. Myers, P. Berry, J. Blythe, *et al.*, “An Intelligent Personal Assistant for Task and Time Management”, *AI Magazine*, vol. 28, no. 2, pp. 47–47, 2007.

[120] K. L. Myers and N. Yorke-Smith, “Proactivity in an Intentionally Helpful Personal Assistive Agent”, in *AAAI Spring Symposium: Intentions in Intelligent Systems*, AAAI Press, 2007, pp. 34–37.

[121] M. A. Neerincx, J. van der Waa, F. Kaptein, and J. van Diggelen, “Using Perceptual and Cognitive Explanations for Enhanced Human-agent Team Performance”, in *Proceedings of the International Conference on Engineering Psychology and Cognitive Ergonomics*, Springer, 2018, pp. 204–214.

[122] B. Pabjan, “Measuring the Social Relations: Social Distance in Social Structure - A Study of Prison Community”, *Acta Physica Polonica Series B*, vol. 36, no. 8, pp. 2559–2574, 2005.

[123] A. Padovitz, S. W. Loke, and A. Zaslavsky, “Towards a Theory of Context Spaces”, in *Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops*, IEEE, 2004, pp. 38–42.

[124] M. Pantic and A. Vinciarelli, "Social Signal Processing", *The Oxford Handbook of Affective Computing*, pp. 84–98, 2014.

[125] R. E. Park and E. W. Burgess, *Introduction to the Science of Sociology*. University of Chicago Press Chicago, 1921.

[126] S. Parrigon, S. E. Woo, L. Tay, and T. Wang, "CAPTION-ing the Situation: A Lexically-derived Taxonomy of Psychological Situation Characteristics.", *Journal of Personality and Social Psychology*, vol. 112, no. 4, pp. 642–681, 2017.

[127] P. Pasotti, C. M. Jonker, and M. B. Van Riemsdijk, "Action Identification Hierarchies for Behaviour Support Agents", in *Proceedings of the Third Workshop on Cognitive Knowledge Acquisition and Applications*, 2017.

[128] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine Learning in Python", *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[129] G. Pereira, R. Prada, and P. A. Santos, "Integrating Social Power into the Decision-making of Cognitive Agents", *Artificial Intelligence*, vol. 241, pp. 1–44, 2016.

[130] L. A. Pervin, "A Free-response Description Approach to the Analysis of Person-Situation Interaction", *ETS Research Bulletin Series*, vol. 1975, no. 2, pp. i–26, 1975.

[131] S. L. Phillips and C. S. Fischer, "Measuring Social Support Networks in General Populations", *Stressful Life Events and Their Contexts*, pp. 223–233, 1981.

[132] C. Pinder, J. Vermeulen, B. R. Cowan, and R. Beale, "Digital Behaviour Change Interventions to Break and Form Habits", *ACM Transactions on Computer-Human Interaction*, vol. 25, no. 3, p. 15, 2018.

[133] R. Polikar, "Ensemble Based Systems in Decision Making", *IEEE Circuits and Systems Magazine*, vol. 6, no. 3, pp. 21–45, 2006.

[134] L. Pu, W. Moyle, C. Jones, and M. Todorovic, "The Effectiveness of Social Robots for Older Adults: A Systematic Review and Meta-analysis of Randomized Controlled Studies", *The Gerontologist*, vol. 59, no. 1, e37–e51, 2019.

[135] J. F. Rauthmann, D. Gallardo-Pujol, E. M. Guillaume, *et al.*, "The Situational Eight DIAMONDS: A Taxonomy of Major Dimensions of Situation Characteristics", *Journal of Personality and Social Psychology*, vol. 107, no. 4, pp. 677–718, 2014.

[136] J. F. Rauthmann and R. A. Sherman, "Measuring the Situational Eight DIAMONDS Characteristics of Situations: An Optimization of the RSQ-8 to the S8*", *European Journal of Psychological Assessment*, vol. 32, no. 2, pp. 155–164, 2016.

[137] A. Reckwitz, "Toward a Theory of Social practices: A Development in Culturalist Theorizing", *European Journal of Social Theory*, vol. 5, no. 2, pp. 243–263, 2002.

[138] M. Ribera and A. Lapedriza, "Can We Do Better Explanations? a Proposal of User-centered Explainable AI", in *Joint Proceedings of the ACM IUI 2019 Workshops*, ACM, vol. 2327, 2019, pp. 38–45.

[139] M. O. Riedl, "Human-centered Artificial Intelligence and Machine Learning", *Human Behavior and Emerging Technologies*, vol. 1, no. 1, pp. 33–36, 2019.

[140] P. S. Ring and A. H. Van de Ven, "Developmental Processes of Cooperative Interorganizational Relationships", *Academy of Management Review*, vol. 19, no. 1, pp. 90–118, 1994.

[141] T. Rist and M. Schmitt, "Applying Socio-psychological Concepts of Cognitive Consistency to Negotiation Dialog Scenarios with Embodied Conversational Characters", *Animating Expressive Characters for Social Interaction*, pp. 213–234, 2008.

[142] M. Riveiro and S. Thill, ""that's (not) the Output I Expected!" on the Role of End User Expectations in Creating Explanations of AI Systems", *Artificial Intelligence*, vol. 298, p. 103 507, 2021.

[143] M. Rokeach, *The Nature of Human Values*. Free Press, 1973.

[144] A. Rosatelli, E. Gedik, and H. Hung, "Detecting F-formations & Roles in Crowded Social Scenes with Wearables: Combining Proxemics & Dynamics using LSTMs", in *8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos*, IEEE, 2019, pp. 147–153.

[145] A. Rosenfeld and S. Kraus, "Predicting Human Decision-making: From Prediction to Action", *Synthesis Lectures on Artificial Intelligence and Machine Learning*, vol. 12, no. 1, pp. 1–150, 2018.

[146] A. Roshchina, J. Cardiff, and P. Rosso, "Twin: Personality-based Intelligent Recommender System", *Journal of Intelligent & Fuzzy Systems*, vol. 28, no. 5, pp. 2059–2071, 2015.

[147] O. Sagi and L. Rokach, "Ensemble Learning: A Survey", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, e1249, 2018.

[148] L. Sanneman and J. A. Shah, "A Situation Awareness-based Framework for Design and Evaluation of Explainable AI", in *International Workshop on Explainable, Transparent Autonomous Agents and Multi-Agent Systems*, Springer, 2020, pp. 94–110.

[149] G. Saucier, T. Bel-Bahar, and C. Fernandez, "What Modifies the Expression of Personality Tendencies? Defining Basic Domains of Situation Variables", *Journal of Personality*, vol. 75, no. 3, pp. 479–504, 2007.

[150] S. H. Schwartz, "Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries", *Advances in Experimental Social Psychology*, vol. 25, no. 1, pp. 1–65, 1992.

[151] ——, "A Proposal for Measuring Value Orientations Across Nations", in *Questionnaire Development Package of the European Social Survey*. 2003, pp. 259–319.

[152] ——, "An Overview of the Schwartz Theory of Basic Values", *Online Readings in Psychology and Culture*, vol. 2, no. 1, 2012.

[153] A. C. Scott, W. J. Clancey, R. Davis, and E. H. Shortliffe, "Explanation Capabilities of Production-based Consultation Systems", University of Stanford, Tech. Rep., 1977.

[154] R. A. Sherman, C. S. Nave, and D. C. Funder, "Properties of Persons and Situations Related to Overall and Distinctive Personality-Behavior Congruence", *Journal of Research in Personality*, vol. 46, no. 1, pp. 87–101, 2012.

[155] M. P. Singh, "Norms as a Basis for Governing Sociotechnical Systems", *ACM Transactions on Intelligent Systems and Technology*, vol. 5, no. 1, pp. 1–23, 2013.

[156] N. A. Stanton, P. M. Salmon, G. H. Walker, E. Salas, and P. A. Hancock, "State-of-science: Situation Awareness in Individuals, Teams and Systems", *Ergonomics*, vol. 60, no. 4, pp. 449–466, 2017.

[157] G. M. Sullivan and A. R. Artino Jr, "Analyzing and Interpreting Data from Likert-type Scales", *Journal of Graduate Medical Education*, vol. 5, no. 4, pp. 541–542, 2013.

[158] M. Tambe, E. Bowring, J. P. Pearce, P. Varakantham, P. Scerri, and D. V. Pynadeth, "Electric Elves: What Went Wrong and Why", *AI Magazine*, vol. 29, no. 2, pp. 23–31, 2008.

[159] T. Therneau and B. Atkinson, *Rpart: Recursive partitioning and regression trees*, R package version 4.1-13, 2018. [Online]. Available: https://CRAN.R-project.org/package=rpart.

[160] M. Tielman, W.-P. Brinkman, and M. A. Neerincx, "Design Guidelines for a Virtual Coach for Post-traumatic Stress Disorder Patients", in *International Conference on Intelligent Virtual Agents*, Springer, 2014, pp. 434–437.

[161] M. L. Tielman, C. M. Jonker, and M. B. van Riemsdijk, "What Should I Do? Deriving Norms from Actions, Values and Context", in *Proceedings of the 10th International Workshop on Modelling and Reasoning in Context*, CEUR Workshop Proceedings, 2018, pp. 35–40.

[162] N. Tintarev and J. Masthoff, "Evaluating the Effectiveness of Explanations for Recommender Systems", *User Modeling and User-Adapted Interaction*, vol. 22, no. 4-5, pp. 399–439, 2012.

[163] ——, "Explaining Recommendations: Design and Evaluation", in *Recommender Systems Handbook*, Springer, 2015, pp. 353–382.

[164] B. Uzzi, "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness", *Administrative Science Quarterly*, vol. 42, no. 1, pp. 35–67, 1997.

[165] M. Valstar, T. Baur, A. Cafaro, *et al.*, "Ask Alice: An Artificial Retrieval of Information Agent", in *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, ACM, 2016, pp. 419–420.

[166] M. B. Van Riemsdijk, C. M. Jonker, and V. Lesser, "Creating Socially Adaptive Electronic Partners: Interaction, Reasoning and Ethical Challenges", in *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems*, International Foundation for Autonomous Agents and Multiagent Systems, 2015, pp. 1201–1206.

[167] J. D. Vargas Quiros, O. Kapcak, H. Hung, and L. Cabrera-Quiros, "Individual and Joint Body Movement Assessed by Wearable Sensing as a Predictor of Attraction in Speed Dates", *IEEE Transactions on Affective Computing*, 2021.

[168] N. Vasilyeva, D. A. Wilkenfeld, and T. Lombrozo, "Goals Affect the Perceived Quality of Explanations", in *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, Cognitive Science Society, 2015, pp. 2469–2474.

[169] J. van der Waa, E. Nieuwburg, A. Cremers, and M. Neerincx, "Evaluating XAI: A Comparison of Rule-based and Example-based Explanations", *Artificial Intelligence*, vol. 291, p. 103 404, 2021.

[170] D. Wang, Q. Yang, A. Abdul, and B. Y. Lim, "Designing Theory-driven User-centric Explainable AI", in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ACM, 2019, pp. 1–15.

[171] O. E. Williamson, *Markets and Hierarchies: Analysis and Antitrust Implications*. New York: Free Press, 1975.

[172] Z. Xiao, M. X. Zhou, Q. V. Liao, *et al.*, "Tell Me About Yourself: Using an AI-powered Chatbot to Conduct Conversational Surveys with Open-ended Questions", *ACM Transactions on Computer-Human Interaction*, vol. 27, no. 3, pp. 1–37, 2020.

[173] Y. Yang, S. J. Read, and L. C. Miller, "The Concept of Situations", *Social and Personality Psychology Compass*, vol. 3, no. 6, pp. 1018–1037, 2009.

[174] S. S. Yau and J. Liu, "Hierarchical Situation Modeling and Reasoning for Pervasive Computing", in *The Fourth IEEE Workshop on Software Technologies for Future Embedded and Ubiquitous Systems*, IEEE, 2006, pp. 5–10.

[175] J. Ye, S. Dobson, and S. McKeever, "Situation Identification Techniques in Pervasive Computing: A Review", *Pervasive and Mobile Computing*, vol. 8, no. 1, pp. 36–66, 2012.

[176] P. Yolum and M. P. Singh, "Service Graphs for Building Trust", in *OTM Confederated International Conferences: "On the Move to Meaningful Internet Systems"*, Springer, 2004, pp. 509–525.

[177] L. Zavala, P. K. Murukannaiah, N. Poosamani, *et al.*, "Platys: From Position to Place-oriented Mobile Computing", *AI Magazine*, vol. 36, no. 2, pp. 50–62, 2015.

[178] R. Zhao, A. Papangelis, and J. Cassell, "Towards a Dyadic Computational Model of Rapport Management for Human-Virtual Agent Interaction", in *International Conference on Intelligent Virtual Agents*, Springer, 2014, pp. 514–527.

[179] M. Ziegler, K. T. Horstmann, and J. Ziegler, "Personality in Situations: Going Beyond the OCEAN and Introducing the Situation Five", *Psychological Assessment*, vol. 31, no. 4, pp. 567–580, 2019.

ACKNOWLEDGEMENTS

In this last chapter I want to take the opportunity to express my gratitude towards everyone who made my PhD journey easier and more enjoyable. First, to my supervisors for their constant trust and support. To Birna, for giving me the freedom to shape this research, and for pushing me to make sure that every chapter is as clear as possible. To Catholijn, for the interactive article review meetings, and for offering advice and giving me the possibility to be flexible when things got difficult. Thank you also to the people with whom I collaborated during these years: to Myrthe, for the help with the experimental designs, and to Pradeep, for the advice during the paper revision stages.

Special thanks to the independent committee members: Prof.dr. M. A. Neerincx, Prof.dr. D. K. J. Heylen, Prof.dr. P. Yolum, Dr. J. Cassens, and Prof.dr. M. E. Warnier. Thank you for reading and appreciating this work, and for giving me the chance to discuss it.

This thesis was part of the CoreSAEP project, so I would like to thank Malte, Myrthe (again) and Sandy for the many interesting discussions. During the course of my PhD the project meetings transformed into the Values-Mental Models bi-weekly meetings, so I would like to thank Luciano, Pradeep, Siddharth, Enrico, Pei-Yu, Ruben and Sietze for providing a platform to share and discuss ideas.

Next, I would like to thank everyone at the Interactive Intelligence group, which is the perfect place to do research while having a great time. Thank you to Elie, Mike and my friend Vincent for being great office mates, to Fran for bringing a social spirit to the group and for the Antwerp tour, to Elena for the chats in italian, to Rolf for the ping pong matches, and to Miguel for always being up for doing something fun and for being a great friend. Thank you to Anita, Ruud and Bart for always being available to help. Also thank you for all the casual chats and hangouts to Aleksander, Bernd, Ding, Frank, Thomas, Rifca, Jinke, Willem-Paul, Frans, Merijn, Carolina, Mani, Davide, Emma, Masha, Jasper, Nele, Mo, Amir, Zuzanna, Wouter, and the previously mentioned people.

Furthermore, I am greatly thankful to all my friends in Albania, the Netherlands and all around Europe. To Ralvi and Thimjo - we started our PhDs in similar times, and discussing the ideas and progress always helped. Thanks to both of you for moving to the Netherlands at some point, and for all the fun times during these years. Thank you to Gert for the fun summer festivals and for visiting in Delft, to Ijon for the fantasy football/basketball challenges, and to Elisa for always checking in and being supportive.

Finally, the biggest thank you goes to my family for their unconditional love and support. Without them, I would have not managed to succeed in this journey. Babi, faleminderit që më ke mësuar që çdo gjë duhet bërë me pasion. Mami, faleminderit që më ke formuar në çdo hap, çdo arritje ka qenë falë teje. Ina, faleminderit që shërben gjithnjë si shtysë për të arritur më shumë, megjithëse pa e kuptuar. Nina, we met at a pivotal moment of my PhD (and life), and I feel extremely lucky for that. Thank you for moving with me to the Netherlands, for always believing in me, and for teaching me the importance of enjoying life. Can't wait for all the exciting adventures ahead.

Curriculum Vitæ

Ilir KOLA

27-10-1992	Born in Peshkopi, Albania.

Education

2014–2017	Master Computer Science University of Freiburg, Germany
2011–2014	Bachelor Computer Engineering Polytechnic University of Tirana, Albania

Work experience

2017–2022	PhD candidate Interactive Intelligence Delft University of Technology, The Netherlands *Thesis:* Enabling Social Situation Awareness in Support Agents *Promotor:* Prof. Dr. C.M. Jonker *Co-Promotor:* Dr. M.B. van Riemsdijk
2015–2017	Research Assistant Cognitive Computation Lab, University of Freiburg, Germany
04–07.2017	Research Intern Max Planck Institute for Human Development, Germany
04–08.2016	Teaching Assistant Machine Learning Lab, University of Freiburg, Germany

List of Publications

6. **I. Kola, R. Isufaj, C. M. Jonker**, *Does Personalization Help? Predicting How Social Situations Affect Personal Values*, Proceedings of the International Conference on Hybrid Human- Artificial Intelligence (2022).

5. **I. Kola, P. K. Murukannaiah, C. M. Jonker, M. B. van Riemsdijk**, *Towards Social Situation Awareness in Support Agents*, IEEE Intelligent Systems (2022).

4. **I. Kola, C. M. Jonker, M. B. van Riemsdijk**, *Using Psychological Characteristics of Situations for Social Situation Comprehension in Support Agents*, Under review (2021).

3. **I. Kola, M. L. Tielman, C. M. Jonker, M. B. van Riemsdijk**, *Predicting the Priority of Social Situations for Personal Assistant Agents*, Proceedings of the International Conference on Principles and Practice of Multi-Agent Systems (2021).

2. **I. Kola, C. M. Jonker, M. L. Tielman, M. B. van Riemsdijk**, *Grouping Situations Based on their Psychological Characteristics Gives Insight into Personal Values*, Proceedings of the Eleventh International Workshop on Modelling and Reasoning in Context (2020).

1. **I. Kola, C. M. Jonker, M. B. van Riemsdijk**, *Who's That? - Social Situation Awareness for Behaviour Support Agents*, Proceedings of the International Workshop on Engineering Multi-Agent Systems (2020).

SIKS Dissertation Series

2016	01	Syed Saiden Abbas (RUN), Recognition of Shapes by Humans and Machines
	02	Michiel Christiaan Meulendijk (UU), Optimizing medication reviews through decision support: prescribing a better pill to swallow
	03	Maya Sappelli (RUN), Knowledge Work in Context: User Centered Knowledge Worker Support
	04	Laurens Rietveld (VU), Publishing and Consuming Linked Data
	05	Evgeny Sherkhonov (UVA), Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers
	06	Michel Wilson (TUD), Robust scheduling in an uncertain environment
	07	Jeroen de Man (VU), Measuring and modeling negative emotions for virtual training
	08	Matje van de Camp (TiU), A Link to the Past: Constructing Historical Social Networks from Unstructured Data
	09	Archana Nottamkandath (VU), Trusting Crowdsourced Information on Cultural Artefacts
	10	George Karafotias (VUA), Parameter Control for Evolutionary Algorithms
	11	Anne Schuth (UVA), Search Engines that Learn from Their Users
	12	Max Knobbout (UU), Logics for Modelling and Verifying Normative Multi-Agent Systems
	13	Nana Baah Gyan (VU), The Web, Speech Technologies and Rural Development in West Africa - An ICT4D Approach
	14	Ravi Khadka (UU), Revisiting Legacy Software System Modernization
	15	Steffen Michels (RUN), Hybrid Probabilistic Logics - Theoretical Aspects, Algorithms and Experiments
	16	Guangliang Li (UVA), Socially Intelligent Autonomous Agents that Learn from Human Reward
	17	Berend Weel (VU), Towards Embodied Evolution of Robot Organisms
	18	Albert Meroño Peñuela (VU), Refining Statistical Data on the Web
	19	Julia Efremova (Tu/e), Mining Social Structures from Genealogical Data
	20	Daan Odijk (UVA), Context & Semantics in News & Web Search
	21	Alejandro Moreno Célleri (UT), From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground
	22	Grace Lewis (VU), Software Architecture Strategies for Cyber-Foraging Systems
	23	Fei Cai (UVA), Query Auto Completion in Information Retrieval
	24	Brend Wanders (UT), Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach

25 Julia Kiseleva (TU/e), Using Contextual Information to Understand Searching and Browsing Behavior
26 Dilhan Thilakarathne (VU), In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains
27 Wen Li (TUD), Understanding Geo-spatial Information on Social Media
28 Mingxin Zhang (TUD), Large-scale Agent-based Social Simulation - A study on epidemic prediction and control
29 Nicolas Höning (TUD), Peak reduction in decentralised electricity systems - Markets and prices for flexible planning
30 Ruud Mattheij (UvT), The Eyes Have It
31 Mohammad Khelghati (UT), Deep web content monitoring
32 Eelco Vriezekolk (UT), Assessing Telecommunication Service Availability Risks for Crisis Organisations
33 Peter Bloem (UVA), Single Sample Statistics, exercises in learning from just one example
34 Dennis Schunselaar (TUE), Configurable Process Trees: Elicitation, Analysis, and Enactment
35 Zhaochun Ren (UVA), Monitoring Social Media: Summarization, Classification and Recommendation
36 Daphne Karreman (UT), Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies
37 Giovanni Sileno (UvA), Aligning Law and Action - a conceptual and computational inquiry
38 Andrea Minuto (UT), Materials that Matter - Smart Materials meet Art & Interaction Design
39 Merijn Bruijnes (UT), Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect
40 Christian Detweiler (TUD), Accounting for Values in Design
41 Thomas King (TUD), Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance
42 Spyros Martzoukos (UVA), Combinatorial and Compositional Aspects of Bilingual Aligned Corpora
43 Saskia Koldijk (RUN), Context-Aware Support for Stress Self-Management: From Theory to Practice
44 Thibault Sellam (UVA), Automatic Assistants for Database Exploration
45 Bram van de Laar (UT), Experiencing Brain-Computer Interface Control
46 Jorge Gallego Perez (UT), Robots to Make you Happy
47 Christina Weber (UL), Real-time foresight - Preparedness for dynamic innovation networks
48 Tanja Buttler (TUD), Collecting Lessons Learned
49 Gleb Polevoy (TUD), Participation and Interaction in Projects. A Game-Theoretic Analysis
50 Yan Wang (UVT), The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains

2017 01 Jan-Jaap Oerlemans (UL), Investigating Cybercrime
02 Sjoerd Timmer (UU), Designing and Understanding Forensic Bayesian Networks using Argumentation
03 Daniël Harold Telgen (UU), Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines
04 Mrunal Gawade (CWI), Multi-core Parallelism in a Column-store
05 Mahdieh Shadi (UVA), Collaboration Behavior
06 Damir Vandic (EUR), Intelligent Information Systems for Web Product Search
07 Roel Bertens (UU), Insight in Information: from Abstract to Anomaly
08 Rob Konijn (VU) , Detecting Interesting Differences:Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery
09 Dong Nguyen (UT), Text as Social and Cultural Data: A Computational Perspective on Variation in Text
10 Robby van Delden (UT), (Steering) Interactive Play Behavior
11 Florian Kunneman (RUN), Modelling patterns of time and emotion in Twitter #anticipointment
12 Sander Leemans (TUE), Robust Process Mining with Guarantees
13 Gijs Huisman (UT), Social Touch Technology - Extending the reach of social touch through haptic technology
14 Shoshannah Tekofsky (UvT), You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior
15 Peter Berck (RUN), Memory-Based Text Correction
16 Aleksandr Chuklin (UVA), Understanding and Modeling Users of Modern Search Engines
17 Daniel Dimov (UL), Crowdsourced Online Dispute Resolution
18 Ridho Reinanda (UVA), Entity Associations for Search
19 Jeroen Vuurens (UT), Proximity of Terms, Texts and Semantic Vectors in Information Retrieval
20 Mohammadbashir Sedighi (TUD), Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility
21 Jeroen Linssen (UT), Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)
22 Sara Magliacane (VU), Logics for causal inference under uncertainty
23 David Graus (UVA), Entities of Interest — Discovery in Digital Traces
24 Chang Wang (TUD), Use of Affordances for Efficient Robot Learning
25 Veruska Zamborlini (VU), Knowledge Representation for Clinical Guidelines, with applications to Multimorbidity Analysis and Literature Search
26 Merel Jung (UT), Socially intelligent robots that understand and respond to human touch
27 Michiel Joosse (UT), Investigating Positioning and Gaze Behaviors of Social Robots: People's Preferences, Perceptions and Behaviors
28 John Klein (VU), Architecture Practices for Complex Contexts

29 Adel Alhuraibi (UvT), From IT-BusinessStrategic Alignment to Performance: A Moderated Mediation Model of Social Innovation, and Enterprise Governance of IT"
30 Wilma Latuny (UvT), The Power of Facial Expressions
31 Ben Ruijl (UL), Advances in computational methods for QFT calculations
32 Thaer Samar (RUN), Access to and Retrievability of Content in Web Archives
33 Brigit van Loggem (OU), Towards a Design Rationale for Software Documentation: A Model of Computer-Mediated Activity
34 Maren Scheffel (OU), The Evaluation Framework for Learning Analytics
35 Martine de Vos (VU), Interpreting natural science spreadsheets
36 Yuanhao Guo (UL), Shape Analysis for Phenotype Characterisation from High-throughput Imaging
37 Alejandro Montes Garcia (TUE), WiBAF: A Within Browser Adaptation Framework that Enables Control over Privacy
38 Alex Kayal (TUD), Normative Social Applications
39 Sara Ahmadi (RUN), Exploiting properties of the human auditory system and compressive sensing methods to increase noise robustness in ASR
40 Altaf Hussain Abro (VUA), Steer your Mind: Computational Exploration of Human Control in Relation to Emotions, Desires and Social Support For applications in human-aware support systems
41 Adnan Manzoor (VUA), Minding a Healthy Lifestyle: An Exploration of Mental Processes and a Smart Environment to Provide Support for a Healthy Lifestyle
42 Elena Sokolova (RUN), Causal discovery from mixed and missing data with applications on ADHD datasets
43 Maaike de Boer (RUN), Semantic Mapping in Video Retrieval
44 Garm Lucassen (UU), Understanding User Stories - Computational Linguistics in Agile Requirements Engineering
45 Bas Testerink (UU), Decentralized Runtime Norm Enforcement
46 Jan Schneider (OU), Sensor-based Learning Support
47 Jie Yang (TUD), Crowd Knowledge Creation Acceleration
48 Angel Suarez (OU), Collaborative inquiry-based learning

2018 01 Han van der Aa (VUA), Comparing and Aligning Process Representations
02 Felix Mannhardt (TUE), Multi-perspective Process Mining
03 Steven Bosems (UT), Causal Models For Well-Being: Knowledge Modeling, Model-Driven Development of Context-Aware Applications, and Behavior Prediction
04 Jordan Janeiro (TUD), Flexible Coordination Support for Diagnosis Teams in Data-Centric Engineering Tasks
05 Hugo Huurdeman (UVA), Supporting the Complex Dynamics of the Information Seeking Process
06 Dan Ionita (UT), Model-Driven Information Security Risk Assessment of Socio-Technical Systems
07 Jieting Luo (UU), A formal account of opportunism in multi-agent systems
08 Rick Smetsers (RUN), Advances in Model Learning for Software Systems

09 Xu Xie (TUD), Data Assimilation in Discrete Event Simulations
10 Julienka Mollee (VUA), Moving forward: supporting physical activity behavior change through intelligent technology
11 Mahdi Sargolzaei (UVA), Enabling Framework for Service-oriented Collaborative Networks
12 Xixi Lu (TUE), Using behavioral context in process mining
13 Seyed Amin Tabatabaei (VUA), Computing a Sustainable Future
14 Bart Joosten (UVT), Detecting Social Signals with Spatiotemporal Gabor Filters
15 Naser Davarzani (UM), Biomarker discovery in heart failure
16 Jaebok Kim (UT), Automatic recognition of engagement and emotion in a group of children
17 Jianpeng Zhang (TUE), On Graph Sample Clustering
18 Henriette Nakad (UL), De Notaris en Private Rechtspraak
19 Minh Duc Pham (VUA), Emergent relational schemas for RDF
20 Manxia Liu (RUN), Time and Bayesian Networks
21 Aad Slootmaker (OUN), EMERGO: a generic platform for authoring and playing scenario-based serious games
22 Eric Fernandes de Mello Araujo (VUA), Contagious: Modeling the Spread of Behaviours, Perceptions and Emotions in Social Networks
23 Kim Schouten (EUR), Semantics-driven Aspect-Based Sentiment Analysis
24 Jered Vroon (UT), Responsive Social Positioning Behaviour for Semi-Autonomous Telepresence Robots
25 Riste Gligorov (VUA), Serious Games in Audio-Visual Collections
26 Roelof Anne Jelle de Vries (UT),Theory-Based and Tailor-Made: Motivational Messages for Behavior Change Technology
27 Maikel Leemans (TUE), Hierarchical Process Mining for Scalable Software Analysis
28 Christian Willemse (UT), Social Touch Technologies: How they feel and how they make you feel
29 Yu Gu (UVT), Emotion Recognition from Mandarin Speech
30 Wouter Beek, The "K" in "semantic web" stands for "knowledge": scaling semantics to the web

---

2019 01 Rob van Eijk (UL),Web privacy measurement in real-time bidding systems. A graph-based approach to RTB system classification
02 Emmanuelle Beauxis Aussalet (CWI, UU), Statistics and Visualizations for Assessing Class Size Uncertainty
03 Eduardo Gonzalez Lopez de Murillas (TUE), Process Mining on Databases: Extracting Event Data from Real Life Data Sources
04 Ridho Rahmadi (RUN), Finding stable causal structures from clinical data
05 Sebastiaan van Zelst (TUE), Process Mining with Streaming Data
06 Chris Dijkshoorn (VU), Nichesourcing for Improving Access to Linked Cultural Heritage Datasets
07 Soude Fazeli (TUD), Recommender Systems in Social Learning Platforms

08 Frits de Nijs (TUD), Resource-constrained Multi-agent Markov Decision Processes
09 Fahimeh Alizadeh Moghaddam (UVA), Self-adaptation for energy efficiency in software systems
10 Qing Chuan Ye (EUR), Multi-objective Optimization Methods for Allocation and Prediction
11 Yue Zhao (TUD), Learning Analytics Technology to Understand Learner Behavioral Engagement in MOOCs
12 Jacqueline Heinerman (VU), Better Together
13 Guanliang Chen (TUD), MOOC Analytics: Learner Modeling and Content Generation
14 Daniel Davis (TUD), Large-Scale Learning Analytics: Modeling Learner Behavior & Improving Learning Outcomes in Massive Open Online Courses
15 Erwin Walraven (TUD), Planning under Uncertainty in Constrained and Partially Observable Environments
16 Guangming Li (TUE), Process Mining based on Object-Centric Behavioral Constraint (OCBC) Models
17 Ali Hurriyetoglu (RUN),Extracting actionable information from microtexts
18 Gerard Wagenaar (UU), Artefacts in Agile Team Communication
19 Vincent Koeman (TUD), Tools for Developing Cognitive Agents
20 Chide Groenouwe (UU), Fostering technically augmented human collective intelligence
21 Cong Liu (TUE), Software Data Analytics: Architectural Model Discovery and Design Pattern Detection
22 Martin van den Berg (VU),Improving IT Decisions with Enterprise Architecture
23 Qin Liu (TUD), Intelligent Control Systems: Learning, Interpreting, Verification
24 Anca Dumitrache (VU), Truth in Disagreement - Crowdsourcing Labeled Data for Natural Language Processing
25 Emiel van Miltenburg (VU), Pragmatic factors in (automatic) image description
26 Prince Singh (UT), An Integration Platform for Synchromodal Transport
27 Alessandra Antonaci (OUN), The Gamification Design Process applied to (Massive) Open Online Courses
28 Esther Kuindersma (UL), Cleared for take-off: Game-based learning to prepare airline pilots for critical situations
29 Daniel Formolo (VU), Using virtual agents for simulation and training of social skills in safety-critical circumstances
30 Vahid Yazdanpanah (UT), Multiagent Industrial Symbiosis Systems
31 Milan Jelisavcic (VU), Alive and Kicking: Baby Steps in Robotics
32 Chiara Sironi (UM), Monte-Carlo Tree Search for Artificial General Intelligence in Games
33 Anil Yaman (TUE), Evolution of Biologically Inspired Learning in Artificial Neural Networks

	34	Negar Ahmadi (TUE), EEG Microstate and Functional Brain Network Features for Classification of Epilepsy and PNES
	35	Lisa Facey-Shaw (OUN), Gamification with digital badges in learning programming
	36	Kevin Ackermans (OUN), Designing Video-Enhanced Rubrics to Master Complex Skills
	37	Jian Fang (TUD), Database Acceleration on FPGAs
	38	Akos Kadar (OUN), Learning visually grounded and multilingual representations
2020	01	Armon Toubman (UL), Calculated Moves: Generating Air Combat Behaviour
	02	Marcos de Paula Bueno (UL), Unraveling Temporal Processes using Probabilistic Graphical Models
	03	Mostafa Deghani (UvA), Learning with Imperfect Supervision for Language Understanding
	04	Maarten van Gompel (RUN), Context as Linguistic Bridges
	05	Yulong Pei (TUE), On local and global structure mining
	06	Preethu Rose Anish (UT), Stimulation Architectural Thinking during Requirements Elicitation - An Approach and Tool Support
	07	Wim van der Vegt (OUN), Towards a software architecture for reusable game components
	08	Ali Mirsoleimani (UL),Structured Parallel Programming for Monte Carlo Tree Search
	09	Myriam Traub (UU), Measuring Tool Bias and Improving Data Quality for Digital Humanities Research
	10	Alifah Syamsiyah (TUE), In-database Preprocessing for Process Mining
	11	Sepideh Mesbah (TUD), Semantic-Enhanced Training Data Augmentation-Methods for Long-Tail Entity Recognition Models
	12	Ward van Breda (VU), Predictive Modeling in E-Mental Health: Exploring Applicability in Personalised Depression Treatment
	13	Marco Virgolin (CWI), Design and Application of Gene-pool Optimal Mixing Evolutionary Algorithms for Genetic Programming
	14	Mark Raasveldt (CWI/UL), Integrating Analytics with Relational Databases
	15	Konstantinos Georgiadis (OUN), Smart CAT: Machine Learning for Configurable Assessments in Serious Games
	16	Ilona Wilmont (RUN), Cognitive Aspects of Conceptual Modelling
	17	Daniele Di Mitri (OUN), The Multimodal Tutor: Adaptive Feedback from Multimodal Experiences
	18	Georgios Methenitis (TUD), Agent Interactions & Mechanisms in Markets with Uncertainties: Electricity Markets in Renewable Energy Systems
	19	Guido van Capelleveen (UT), Industrial Symbiosis Recommender Systems
	20	Albert Hankel (VU), Embedding Green ICT Maturity in Organisations
	21	Karine da Silva Miras de Araujo (VU), Where is the robot?: Life as it could be

22 Maryam Masoud Khamis (RUN), Understanding complex systems implementation through a modeling approach: the case of e-government in Zanzibar
23 Rianne Conijn (UT), The Keys to Writing: A writing analytics approach to studying writing processes using keystroke logging
24 Lenin da Nobrega Medeiros (VUA/RUN), How are you feeling, human? Towards emotionally supportive chatbots
25 Xin Du (TUE), The Uncertainty in Exceptional Model Mining
26 Krzysztof Leszek Sadowski (UU), GAMBIT: Genetic Algorithm for Model-Based mixed-Integer opTimization
27 Ekaterina Muravyeva (TUD), Personal data and informed consent in an educational context
28 Bibeg Limbu (TUD), Multimodal interaction for deliberate practice: Training complex skills with augmented reality
29 Ioan Gabriel Bucur (RUN), Being Bayesian about Causal Inference
30 Bob Zadok Blok (UL), Creatief, Creatieve, Creatiefst
31 Gongjin Lan (VU), Learning better – From Baby to Better
32 Jason Rhuggenaath (TUE), Revenue management in online markets: pricing and online advertising
33 Rick Gilsing (TUE), Supporting service-dominant business model evaluation in the context of business model innovation
34 Anna Bon (MU), Intervention or Collaboration? Redesigning Information and Communication Technologies for Development
35 Siamak Farshidi (UU), Multi-Criteria Decision-Making in Software Production

---

2021 01 Francisco Xavier Dos Santos Fonseca (TUD),Location-based Games for Social Interaction in Public Space
02 Rijk Mercuur (TUD), Simulating Human Routines:Integrating Social Practice Theory in Agent-Based Models
03 Seyyed Hadi Hashemi (UVA), Modeling Users Interacting with Smart Devices
04 Ioana Jivet (OU), The Dashboard That Loved Me: Designing adaptive learning analytics for self-regulated learning
05 Davide Dell'Anna (UU), Data-Driven Supervision of Autonomous Systems
06 Daniel Davison (UT), "Hey robot, what do you think?" How children learn with a social robot
07 Armel Lefebvre (UU), Research data management for open science
08 Nardie Fanchamps (OU), The Influence of Sense-Reason-Act Programming on Computational Thinking
09 Cristina Zaga (UT), The Design of Robothings. Non-Anthropomorphic and Non-Verbal Robots to Promote Children's Collaboration Through Play
10 Quinten Meertens (UvA), Misclassification Bias in Statistical Learning
11 Anne van Rossum (UL), Nonparametric Bayesian Methods in Robotic Vision
12 Lei Pi (UL), External Knowledge Absorption in Chinese SMEs

13 Bob R. Schadenberg (UT), Robots for Autistic Children: Understanding and Facilitating Predictability for Engagement in Learning
14 Negin Samaeemofrad (UL), Business Incubators: The Impact of Their Support
15 Onat Ege Adali (TU/e), Transformation of Value Propositions into Resource Re-Configurations through the Business Services Paradigm
16 Esam A. H. Ghaleb (UM), Bimodal Emotion Recognition from Audio-Visual Cues
17 Dario Dotti (UM), Human Behavior Understanding from motion and bodily cues using deep neural networks
18 Remi Wieten (UU), Bridging the Gap Between Informal Sense-Making Tools and Formal Systems - Facilitating the Construction of Bayesian Networks and Argumentation Frameworks
19 Roberto Verdecchia (VU), Architectural Technical Debt: Identification and Management
20 Masoud Mansoury (TU/e), Understanding and Mitigating Multi-Sided Exposure Bias in Recommender Systems
21 Pedro Thiago Timbó Holanda (CWI), Progressive Indexes
22 Sihang Qiu (TUD), Conversational Crowdsourcing
23 Hugo Manuel Proença (LIACS), Robust rules for prediction and description
24 Kaijie Zhu (TUE), On Efficient Temporal Subgraph Query Processing
25 Eoin Martino Grua (VUA), The Future of E-Health is Mobile: Combining AI and Self-Adaptation to Create Adaptive E-Health Mobile Applications
26 Benno Kruit (CWI & VUA), Reading the Grid: Extending Knowledge Bases from Human-readable Tables
27 Jelte van Waterschoot (UT), Personalized and Personal Conversations: Designing Agents Who Want to Connect With You
28 Christoph Selig (UL), Understanding the Heterogeneity of Corporate Entrepreneurship Programs

---

2022 1 Judith van Stegeren (UT), Flavor text generation for role-playing video games
2 Paulo da Costa (TU/e), Data-driven Prognostics and Logistics Optimisation: A Deep Learning Journey
3 Ali el Hassouni (VUA), A Model A Day Keeps The Doctor Away: Reinforcement Learning For Personalized Healthcare
4 Ünal Aksu (UU), A Cross-Organizational Process Mining Framework
5 Shiwei Liu (TU/e), Sparse Neural Network Training with In-Time Over-Parameterization
6 Reza Refaei Afshar (TU/e), Machine Learning for Ad Publishers in Real Time Bidding
7 Sambit Praharaj (OU), Measuring the Unmeasurable? Towards Automatic Co-located Collaboration Analytics
8 Maikel L. van Eck (TU/e), Process Mining for Smart Product Design
9 Oana Andreea Inel (VUA), Understanding Events: A Diversity-driven Human-Machine Approach

10 Felipe Moraes Gomes (TUD), Examining the Effectiveness of Collaborative Search Engines
11 Mirjam de Haas (UT), Staying engaged in child-robot interaction, a quantitative approach to studying preschoolers' engagement with robots and tasks during second-language tutoring
12 Guanyi Chen (UU), Computational Generation of Chinese Noun Phrases
13 Xander Wilcke (VUA), Machine Learning on Multimodal Knowledge Graphs: Opportunities, Challenges, and Methods for Learning on Real-World Heterogeneous and Spatially-Oriented Knowledge
14 Michiel Overeem (UU), Evolution of Low-Code Platforms
15 Jelmer Jan Koorn (UU), Work in Process: Unearthing Meaning using Process Mining
16 Pieter Gijsbers (TU/e), Systems for AutoML Research
17 Laura van der Lubbe (VUA), Empowering vulnerable people with serious games and gamification