What should I do? Deriving norms from actions,values and context

Myrthe L. Tielman[1], **Catholijn M. Jonker**[1,2], **M. Birna van Riemsdijk**[1],
[1] Interactive Intelligence Group, TU Delft, The Netherlands
{m.l.tielman, c.m.jonker, m.b.vanriemsdijk}@tudelft.nl
[2] Leiden Institute of Advanced Computer Science, Leiden University
c.m.jonker@liacs.leidenuniv.nl

Abstract

Behavior support technology is increasingly used to assist people in daily life activities. To do this properly, it is important that the technology understands what really motivates people. What values underlie their actions, but also the influence of context, and how this can be translated to norms which govern behavior. In this paper, we expand a framework describing action hierarchies and values to include the role of context. Moreover, we present a method to derive specific norms for behavior from this information on actions, values and context. Behavior support technology can use this framework to reason about peoples ideal behavior, and so better offer personalized assistance.

1 Introduction

Behavior support technology is used increasingly in daily life. We have technology which reminds us of meetings or to take our medicine [Milić *et al.*, 2018] [Zhou *et al.*, 2012], to help us eat healthier [Schoffman *et al.*, 2013], and even with diverse medical problems such as dementia [Carrillo *et al.*, 2009], depression [Karyotaki *et al.*, 2017] or kidney transplants [Wang *et al.*, 2017]. As the role that technology plays in our daily lives increases, it becomes more important that technology supports us in a flexible way [van Riemsdijk *et al.*, 2015]. Ideally, you would want your daily behavior support app to understand it only needs to remind you to take your umbrella when it is going to rain, to remind you of job interviews earlier than meetings with friends, and to call your doctor only if the medicine you forgot was absolutely crucial. You want it to take into account the context you are in, and understand how this would affect your ideal choices.

Although all these flexibilities could be incorporated into technology separately, a more sustainable approach seems to ensure that the technology itself understands our motivations as humans. To realize this, many systems introduce *values* [van de Poel, 2015], [Cranefield *et al.*, 2017], concepts which refer to what a person or group of people consider important in life [Friedman *et al.*, 2006]. Values are used to identify the underlying reasons for our actions, and are particularly suitable for this purpose due to their generalizability and stability over time [Schwartz, 1992].

However, current systems do not always explicitly take into account the role of context. This can be problematic, as a single action might support different values in different ways, depending on the situation. For instance, biking can promote the values health and sustainability. However, when it is snowing, this effects how much health is promoted, as biking through snow is decidedly less healthy. It has no effect on the sustainability of the action though.

This example shows that context cannot be ignored when reasoning about how actions promote values. We therefore propose a framework which does not just include a representation of how values relate to our actions, but which also models the role of context. Our focus in this paper lies on reasoning about context, so not necessarily in modeling context itself, as done, for example, in [Kola *et al.*, 2018].

Aside from understanding our values in context, technology is also required to reason about what this means concretely. We want it to understand what choices are best, in other words: what norms we wish to adhere to, given our values, the actions we can take, and the context we are in. Norms "regulate the interactions between an individual and the society" [Balke *et al.*, 2013], and are often used in agent systems to reason about the behavior of artificial agents [Santos *et al.*, 2017], but can be used similarly to reason about what behavior to support in users [van Riemsdijk *et al.*, 2015]. Most multi-agent systems employing norms either derive them beforehand based on goals, or look at how norms emerge in a society based on what actions agents perform [Balke *et al.*, 2013]. These systems look at how norms govern behavior of groups. However, these approaches are less suitable when considering personal norms. For this reason, [Criado *et al.*, 2013] use a more human-inspired model, but they do not explicitly consider an individual's values. Other work takes a different approach, and looks instead at how values can govern behavior directly [Cranefield *et al.*, 2017]. [Bench-Capon and Modgil, 2017] do consider how norms, actions and values work together, but they employ the societal norms, instead of norms based on the individuals own values. Finally, work from the angle of values-sensitive design considers how norms relate to underlying values, but often do not translate this into specific options for behavior [van de Poel, 2013]. In our framework, we propose to bring actions, values, context and norms together, as shown in Figure 1.

In section 2. we describe the action framework, as well

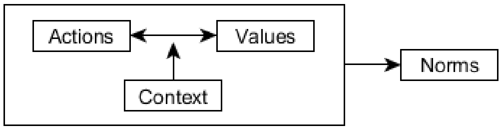

Figure 1: Schematic representation of our framework, including actions, values and context, and norms derived from this information.

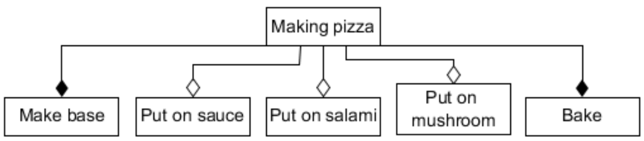

Figure 2: Example of a tree with part-of relationships, describing making pizza. The black diamond links indicate necessity.

as how values relate to these actions. Section 3 discusses the role of context. Finally, in section 4 we describe how specific norms can automatically be derived from this framework. These norms allow a behavior support agent to reason about what the ideal behavior of the user would be.

2 Actions and Values

2.1 Action hierarchies

In order to support people in their daily behavior, it is important to understand how they themselves conceptualize their actions and the relations between them. To this end, Pasotti et al (2017) developed a knowledge representation capable of describing Action Identification Hierarchies (AIH) [Pasotti *et al.*, 2017]. The core concepts of the action framework described in our paper originate in this work.

At the core, AIH describes relationships between actions. For this paper, we only consider *part-of* relations. A part-of relation from action A to action B describes that doing action B is a part of doing the action A. So one can do B while doing A, but doing A entails more than just doing B.

The original AIH's can include a multitude of relations describing full behavior trees. For the framework presented in this paper, however, we need only consider two layers at a time. So our AIH's will only consist of one top action, which has a *part-of* relationship with at least two child actions. Figure 2. shows an example of such a tree.

2.2 Values

The question of how to formalize the relationship between our actions and values is a complex one, which is dealt with in different ways across frameworks. [van der Weide *et al.*, 2009] relate values to state changes, and define how a value can either be demoted or promoted by such a change. A similar approach is taken by [**?**], and by [Pasotti *et al.*, 2016], who link values to postconditions of actions. [Sartor, 2010] has a slightly different approach, adding how much a certain choice affects a value.

For this paper, our focus is on the role of context and norm derivation. Therefore, we will employ a relatively simple formalization of values. However, this could be expanded to be more complex without consequences for the rest of the framework, as long as the following criteria are met.

First, we assume there is a relationship between an action and value which denotes how much this action demotes or promotes the value. Secondly, we assume commensurability in this relationship, so we can explicitly compare how much different actions promote a value. This second assumption is not a trivial one [van de Poel, 2015], but important for the computability of the impact of values for an agent.

For this framework, we propose a simple number which expresses how much an action demotes (negative numbers) or promotes (positive numbers) a value. If no explicit relationship between an action and value is given, we assume the action does not affect the value.

A distinction which is nearly always made in the literature, is between the impact an action has on a value, and the importance an individual gives to a value. For the purpose of this paper we only describe the first relation, and leave individual ordering out of the picture. However, as long as a commensurability of values is maintained, this impact of individual preference could easily be added to the framework.

3 The role of Context

Consider the following scenario:

John has a behavior support agent to help him live healthier, as he has problems with his back. One way it does this, is by encouraging John not to carry around unnecessary things such as an umbrella. However, one day John gets caught in the rain, and catches a cold because he did not have his umbrella with him.

This example illustrates the role context can play. Although in general, not taking an umbrella is good for the value health, this does not hold when it is raining. Ideally, John's behavior support agent understands this, and will advise him to take the umbrella only if it is going to rain.

In this example, context is the type of weather. However, context can be any situational circumstance which is not captured in the definition of the actions themselves. Other examples of circumstances which can affect the value-action relationship are time (of day or year), social situation or location.

In our framework, we define a *contextual factor* as a tuple $\langle s, a, v, w \rangle$, where s is a situational property, which when present affects the numerical relationship between an action a and a value v with weight w. Weight can be both positive or negative. So we have a 'default' relationship between an action and value, and the contextual factor modifies this relationship. For instance, given that the action *taking an umbrella* promotes the value *health* by -1, the context of *rain* might influence this by weight +3. The contextual factor would be defined as $\langle$rain, takingUmbrella, health, +3$\rangle$, and this would mean that when it is raining, taking the umbrella actually promotes health by 2. This means the assumption of commensurability of value-action relationships is extended to the influence of context.

Information about context does not need to be present in our framework. If for an action-value relationship no contextual factors are defined, we simply assume this relation is always the same.

4 Deriving Norms

Norms are often used instead of values because they give explicit rules for behavior. However, norms are not as general or stable as values. For this reason, some frameworks introduce values as underlying motivations for norms [Kayal *et al.*, 2014]. In our framework, we take the opposite approach, and instead propose to define norms based on value-action-context information. This not only gives us a clear insight into the relation between norms and values, but the additional flexibility to automatically change norms if context, actions or values change.

Formally, we define a norm as a tuple $\langle deoc, e, ac, a, C\rangle$, where $deoc$ is the deontic modality, obligation, prohibition or permission. e is the entity, i.e. the person whose actions the tree describes. ac is the parent action, which can be seen as the action context. This describes during which action the norm is relevant. a is the child action, so the behavior the norm describes. Finally, C is a set of all the different situational properties c, for which the norm holds. C can be empty, if no specific context is specified. This representation is inspired by existing normative frameworks, e.g. [Balke *et al.*, 2013], [Singh, 1999].

In our normative system, we do not define norms for 'negative' situatons. For instance, we might define that one needs to take an umbrella when it is raining, but not that one should not take an umbrella when it is not raining. This is done to avoid having to check for the absence of a situational property. Instead, we introduce the rule that *a more specific norm will always overrule a more general norm about the same* e*, ac and a, if all situational properties C in the more specific norm hold*. So a norm A is defined as more specific than B if C in norm A includes more situational properties c than C in norm B, so if $C_B \subset C_A$. If we wish to express that one should take the umbrella only when it is raining, one would have one norm expressing *do not take the umbrella*, and one *take the umbrella when it is raining*. When it is raining, the second norm overrides the first. This rule follows the concept of *lex specialis*, specifying that the more specific norm has priority [Balke *et al.*, 2013].

Figure 3 is a graphical representations of a *part-of* structure, describing the actions and values for a user, and situational properties which are of influence. From Figure 3 , we can derive the following norms:

1. *When preparing for driving, find car keys.*
2. *When preparing for driving, you may take sunglasses.*
3. *When preparing for driving in the bright sun, take sunglasses.*
4. *When preparing for driving and running late, do not take sunglasses.*
5. *When preparing for driving in the bright sunlight and running late, take sunglasses.*

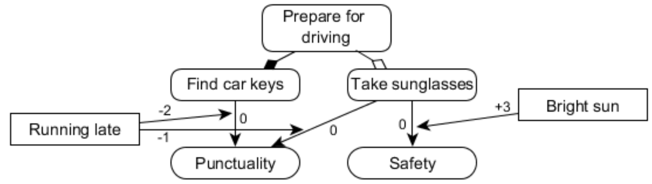

Figure 3: Part-of tree for 'prepare for driving'. With values 'punctuality' and 'safety', and situation 'running late' and 'bright sun'. The black part-of relation indicates a necessity.

For deriving all norms from such a tree, the following procedure is followed, where entity e always refers to the person whose actions the tree describes.

First, obligations can be derived for all *necessary* part-of child actions a of parent action ac. These take the form $\langle obligation, e, ac, a, C\rangle$ where $C = \emptyset$. No other norms are derived for these *necessary* actions. This means that whether these *necessary* actions support the user's values given the situation does not affect the norm, as these actions always need to be included to perform the parent action.

Applying this to Figure 3, we get the formal notation for norm 1:

- $\langle obl, e, prepForDriving, findCarKeys, \emptyset\rangle$ [1]

For all other part-of child actions a of parent action ac, values and context are relevant. We use N_a^i to denote the set of norms derived for action a with set of situational properties C such that $|C| \leq i$. Below we inductively define N_a^i.

1. **Derivation of N_a^0.** Let $s_a = \sum\{v | v \in v(a)\}$, where $v(a)$ is the set of value numericals of a. The norm for a with $|C| = \emptyset$ is $\langle deoc(s_a), e, ac, a, \emptyset\rangle$, where *deoc* depends on the sign of s_a as follows: $deoc(n)$ for any number n is defined to be *obligation* if $n > 0$, $deoc(n)$ = *permission* if $n = 0$, and $deoc(n)$ = *prohibition* if $n < 0$.

   Following this step for Figure 3, we can derive norm 2, formally:

   - $\langle per, e, prepForDriving, sunglasses, \emptyset\rangle$

2. **Derivation of N_a^1.** Next, for each (single) situational property f in the total set present in the tree F, we conditionally decide to add a norm. Let $f \in F$. We define $s_a^f = \sum\{v + n_a^f | v \in v(a)\}$, where n_a^f denotes the numerical value associated to situational property f with respect to action a. If $sign(s_a^f) \neq sign(s_a)$, then the following norm is added: $\langle deoc(s_a^f), e, ac, a, \{f\}\rangle$.

   Following this step for Figure 3, we can derive norms 3 and 4, formally:

   - $\langle obl, e, prepForDriving, sunglasses, \{sun\}\rangle$
   - $\langle pro, e, prepForDriving, sunglasses, \{late\}\rangle$

[1] We abbreviate *obligation* to *obl*, *prohibition* to *pro*, *permission* to *per*, *prepare for driving* to *prepForDriving*, *Running late* to *late*, *Bright sun* to *sun*, *find car keys* to *findCarKeys* and *Take sunglasses* to *sunglasses* in the formal norms.

3. **Derivation of N_a^{i+1}.** Following this, norms are added, depending on whether or not progressively complex combinations of situational properties change the *deoc*. By progressively complex combinations of situational properties we mean that we consider $\mathcal{P}_i(F) = \{t \in \mathcal{P}(F) \mid |t| = i\}$ with cardinality for increasing i, until $i = |F|$. Similarly, $\mathcal{P}_{\leq i}(F) = \{t \in \mathcal{P}(F) \mid |t| \leq i\}$. A norm is added at step $i+1$ only if this larger set of situational properties changes the sign compared to the signs of previously added norms at step i which partly include the same situations.

   More formally, we start with base case is $\mathcal{P}_1(F)$ as described in point 2. Iteratively, when we have calculated N_a^i for $\mathcal{P}_i(F)$, we derive N_a^{i+1} by considering the following for any element $t \in \mathcal{P}_{i+1}(F)$:

   We define $s_a^t = \sum\{v + n_a^t | v \in v(a)\}$, where n_a^t is the sum of context numericals associated with the situational properties in set t with respect to with action a.

   We add a norm $\langle deoc(s_a^t), e, ac, a, t\rangle$ for action a with $t \in \mathcal{P}_{i+1}(F)$, if there is a $t' \in \mathcal{P}_{\leq i}(F)$ such that:

   - $t' \subset t$.
   - There exists a norm $n' \in N_a^i$ of the form $\langle deoc, e, ac, a, t'\rangle$ such that there is no norm $n'' \in N_a^i$ of the form $\langle deoc, e, ac, a, t''\rangle$ where $|t''| > |t'|$.
   - $sign(s_a^t) \neq sign(s_a^{t'})$

   Applied to Figure 3 we can formally derive the final norm, norm 5:

   - $\langle obl, e, prepForDriving, sunglasses, \{sun, late\}\rangle$

5 Discussion

The method for deriving norms from action-value-context information as presented in this paper, generates norms with several specific characteristics which are interesting to note. Firstly, we only consider trees which define what actions are a part of another action. This means all norms describe whether or not to include action A while doing action B. However, [Pasotti *et al.*, 2017] describe another type of relationship between actions, namely *concretisations*. These define action A as a more concrete way of doing action B. One of the directions for future research would be to define how to derive norms from this other type of tree.

Our framework currently does not include any way to derive social norms, for instance where one person has an obligation towards another [Singh, 1999]. This is due to the nature of the action description, which does not have a notion of actions performed with or for someone else. The framework can describe an action such as *sending a text*, but it leaves implicit who it is sent to. This lack of explicit representation of other people means these can also not be made explicit in the norms derived. This also implies that the types of norms we derive are slightly different than used in many multi-agent systems, where norms govern social behavior of agents. Instead, our norms can be seen as personal preferences for an individual's own behavior. Although a social aspect would be a very useful extension of our framework, it makes sense to start with individual behavior in the context of a behavior support agent. After all, this agent will need to understand the behavior and wishes of this particular individual first.

When expressing choices for actions in norms, the observation can be made that some norms somehow seem 'stronger' than others. Using the norms generated from Figure 3, the norm *find the car keys when driving to work* seems of a different order than *do not take sunglasses when running late*. They both make sense, but while ignoring the second norm might just make you a minute late, the first will cause you not to arrive at all. While our action framework partly makes this distinction with the *necessary* part-of link, this is not yet translated into the norm. Moreover, some non-necessary actions might still be more important than others. Some normative frameworks include the notion of *sanction*, which could be used to express the difference between these norms. Another option might be to rank norms based on priority, for instance based on how much an action promotes or demotes values. Our framework, however, as yet does not include a way to express the effect of choosing to include an action.

Another difference between the norms we derive and some other frameworks, is that we include two actions instead of one. This is an indirect result from expressing actions in hierarchies. In a way, what we denote as our *parent* action, is also a form of context, giving our norms two explicit contexts. The first is the *action context*, expressing what the user is doing at the time at a higher level. The other type is the situational context as described in section 3.

With regards to the way in which actions and values relate, we currently assume commensurability. This means we can compare how actions relate to values on a set scale. This assumption is not trivial, however, and further research might therefore shed light on whether this assumption can be relaxed in any way. The key point of our framework is that different actions can be compared based on how well they promote or demote values. Further research might reveal a method to do this without assuming commensurability.

Finally, our framework does not currently include any explicit preference ordering of values. Because the individual differences of value preferences are an important advantage of values, this might be one of the first additions to the framework. One simple method to do this would be to take the weighted sum of values given their ordering. Inspiration could also be taken from [Cranefield *et al.*, 2017], who include a threshold for values. If a value has already reached this threshold, it temporarily becomes less important. In whatever way this is done, as long as one can 'calculate' score for an action given the values it promotes and the context, the method for deriving norms will still work.

6 Conclusion

In this paper, we present a framework which represents hierarchical trees of actions, including how these promote and demote values, and the influence of context. Moreover, we present a method for automatically deriving norms from this information, capable of generating obligations, permissions and prohibitions for behavior. These norms could serve as a starting point for behavior support technology, which could use them to better take into account both the users values and the context they are in while offering support.

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

References

[Balke *et al.*, 2013] Tina Balke, Celia da Costa Pereira, Frank Dignum, Emiliano Lorini, Antonino Rotolo, Wamberto Vasconcelos, and Serena Villata. *Normative Multi-Agent Systems*. Schloss Dagstuhl, 2013.

[Bench-Capon and Modgil, 2017] Trevor Bench-Capon and Sanjay Modgil. Norms and value based reasoning: justifying compliance and violation. *Artificial Intelligence and Law*, 25:29–64, 2017.

[Carrillo *et al.*, 2009] Maria C. Carrillo, Eric Dishman, and Tim Plowman. Everyday technologies for alzheimer's disease care: Research findings, directions, and challenges. *Alzheimer's & Dementia*, 5(6):479 – 488, 2009.

[Cranefield *et al.*, 2017] S. Cranefield, M. Winikoff, V. Dignum, and F. Dignum. No pizza for you: Value-based plan selection in BDI agents. In *International Joint Conference on Artificial Intelligence*, 2017.

[Criado *et al.*, 2013] N. Criado, E. Argente, P. Noriega, and V. Botti. Human-inspired model for norm compliance decision making. *Information Sciences*, 245:218–239, 2013.

[Friedman *et al.*, 2006] Batya Friedman, Peter H. Kahn Jr., and Alan Borning. *Human-Computer Interaction and Management Information Systems: Foundations Advances in Management Information Systems, Volume 5 (Advances in Management Information Systems),*, chapter Value Sensitive Design and Information Systems, pages 348–372. M.E. Sharpe, 2006.

[Karyotaki *et al.*, 2017] E Karyotaki, H Riper, J Twisk, Adriaan Hoogendoorn, Annet Kleiboer, Adriana Mira, Andrew Mackinnon, Bjorn Meyer, Cristina Botella, Elizabeth Littlewood, Gerhard Andersson, Helen Christensen, Jan P. Klein, Johanna Schroder, Juana Breton-Lopez, Justine Scheider, Kathy Griffiths, Louise Farrer, Marcus J. H. Huibers, Rachel Phillips, Simon Gilbody, Steffen Moritz, Thomas Berger, Victor Pop, Viola Spek, and Pim Cuijpers. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: A meta-analysis of individual participant data. *JAMA Psychiatry*, 74(4):351–359, 2017.

[Kayal *et al.*, 2014] Alex Kayal, Willem-Paul Brinkman, Rianne Gouman, Mark A. Neerincx, and M. Birna van Riemsdijk. A value-centric model to ground norms and requirements for epartners of children. In *Coordination, Organizations, Institutions, and Norms in Agent Systems*, 2014.

[Kola *et al.*, 2018] Ilir Kola, Catholijn M. Jonker, and M. Birna van Riemsdijk. Modemodel the social environment: Towards socially adaptive electronic partners. In *MRC - Tenth International Workshop Modelling and Reasoning in Context, Held at FAIM*, 2018. Under revision at the AAMAS/IJCAI Workshop on Modeling and Reasoning in Context.

[Milić *et al.*, 2018] Eleonora Milić, Dragan Janković, and Aleksandar Milenković. Health care domain mobile reminder for taking prescribed medications. In Georgi Stojanov and Andrea Kulakov, editors, *ICT Innovations 2016*, pages 173–181, Cham, 2018. Springer International Publishing.

[Pasotti *et al.*, 2016] Pietro Pasotti, M. Birna van Riemsdijk, and Catholijn M. Jonker. Representing human habits: towards a habit support agent. In *European Conference on Artificial Intelligence*, 2016.

[Pasotti *et al.*, 2017] Pietro Pasotti, Catholijn M. Jonker, and M. Birna van Riemsdijk. Action identification hierarchies for behaviour support agents. In *Workshop on Cognitive Knowledge Acquisition and Applications*, 2017.

[Santos *et al.*, 2017] J.S. Santos, J.O. Zahn, E.A. Silvestre, V.T. Silva, and W.W. Vasconcelos. Detection and resolution of normative conflicts in multi-agent systems: a literature survey. *Journal of Autonomous Agent Multi-Agent Systems*, 31:1236–1282, 2017.

[Sartor, 2010] G. Sartor. Doing justice to rights and values: teleological reasoning and proportionality. *Artif Intell Law*, 18:175–215, 2010.

[Schoffman *et al.*, 2013] Danielle E. Schoffman, Gabrielle Turner-McGrievy, Sonya J. Jones, and Sara Wilcox. Mobile apps for pediatric obesity prevention and treatment, healthy eating, and physical activity promotion: just fun and games? *Translational Behavioral Medicine*, 3(3):320–325, 2013.

[Schwartz, 1992] S.H. Schwartz. Universals in the content and structure of values: theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology*, 25:1–65, 1992.

[Singh, 1999] Munindar P. Singh. An ontology for commitments in multiagent systems: Toward a unification of normative concepts. *Artificial Intelligence and Law*, 7:97–113, 1999.

[van de Poel, 2013] Ibo van de Poel. *Translating Values into Design Requirements*, chapter Philosophy and Engineering: Reflections on Practice, Principles and Process. Springer, 2013.

[van de Poel, 2015] Ibo van de Poel. *Handbook of Ethics, Values and Technological Design*, chapter Conflicting Values in Design for Values, pages 89–115. Springer, 2015.

[van der Weide *et al.*, 2009] T.L. van der Weide, F. Dignum, J.-J. Ch. Meyer, H. Prakken, and G.A.W. Vreeswijk. Practical reasoning using values? Giving meaning to values. In *Proceedings of the 6th international conference on Argumentation in Multi-Agent Systems*, 2009.

[van Riemsdijk *et al.*, 2015] M. Birna van Riemsdijk, Catholijn M. Jonker, and Victor Lesser. Creating socially adaptive electronic partners. In *International Conference on Autonomous Agents and Multiagent Systems*, 2015.

[Wang *et al.*, 2017] Wenxin Wang, Céline L. van Lint, Willem-Paul Brinkman, Ton J. M. Rövekamp, Sandra van Dijk, Paul J. M. van der Boog, and Mark A. Neerincx. Renal transplant patient acceptance of a self-management support system. *BMC Medical Informatics and Decision Making*, 17(1):58, May 2017.

[Zhou *et al.*, 2012] Shandan Zhou, Chao-Hisen Chu, Zhiwen Yu, and Jungyoon Kim. A context-aware reminder system for elders based on fuzzy linguistic approach. *Expert Systems with Applications*, 39(10):9411 – 9419, 2012.