A Computational Approach for Predicting Individuals' Response Patterns in Human Syllogistic Reasoning

Emmanuelle-Anna Dietz Saldanha (emmanuelle.dietz@tu-dresden.de)
International Center for Computational Logic, TU Dresden, Germany

Robert Schambach (robertschambach@mac.com)[1]
International Center for Computational Logic, TU Dresden, Germany

Abstract

One challenge within cognitive psychology on human reasoning is modeling a wide range of tasks within a certain theory. Recently, a meta-study on human syllogistic reasoning has shown that none of the established theories seemed to adequately match the human data. Possible reasons for this sobering result could be that (i) these theories do not account for differences among reasoners and (ii) they presuppose the same assumptions throughout all 64 syllogistic reasoning tasks. In this paper, we will address both aspects by proposing *clustering by principle patterns* for syllogistic reasoning based on cognitive principles, which have their roots in the literature of cognitive science and philosophy of language. These principles determine how the tasks are formally represented within the weak completion semantics, a logic programming approach that has already been successfully applied for modeling various human reasoning episodes. We will develop a generic cognitive characterization of (i) the reasoners and (ii) the tasks by integrating the results of a machine learning algorithm with underlying cognitive principles. These principles provide a cognitively plausible characterization of the response patterns that cover the population of reasoners. *Clustering by principle patterns* achieves the highest prediction accuracy compared to the available benchmark models, and gives insights to the differences among (i) the reasoners and among (ii) the explaining principles throughout the tasks.

Keywords: Cognitive Modeling, Syllogistic Reasoning Task, Individual Reasoning Patterns

Introduction

A wide range of cognitive theories have been proposed in the past (cf. Wetherick and Gilhooly (1995); Chater and Oaksford (1999); Woodworth and Sells (1935); Johnson-Laird (1983); Rips (1994); Chapman and Chapman (1959); Polk and Newell (1995)), aiming to explain the majority of participants' response in reasoning tasks (cf. Wason (1968); Byrne (1989); Khemlani and Johnson-Laird (2012)). Usually, their underlying assumptions plausibly explain by a very particular response pattern within a certain task, namely the responses of the majority of participants (cf. suppression effect in (Byrne, 1989)). It seems that the adequacy of a cognitive theory is confirmed, if these assumptions can explain the majority of participants' responses for various reasoning tasks. However, psychological results have shown that different significant response patterns exist, i.e. there are individual differences among participants, which the theories should account for (Johnson-Laird & Khemlani, 2016; Ragni, Kola, & Johnson-Laird, 2017). Recently a few approaches have been developed that (i) aim at explaining these differences (cf. Johnson-Laird and Khemlani (2016); Breu, Ind, Mertesdorf, and Ragni (2019); Dietz Saldanha and Schambach (2019)), but they did not yet address (ii) the differences among tasks.

Natural Language Sentence	Classical Logic	Mood
all artists are bakers	$\forall X(a(X) \rightarrow b(X))$	Aab
some artists are bakers	$\exists X(a(X) \wedge b(X))$	Iab
no artists are bakers	$\forall X(a(X) \rightarrow \neg b(X))$	Eab
some artists are not bakers	$\exists X(a(X) \wedge \neg b(X))$	Oab

Table 1: The four moods and their formalization.

Figure	Premise 1	Premise 2
1	artists – bakers	bakers – chemists
2	bakers – artists	chemists – bakers
3	artists – bakers	chemists – bakers
4	bakers – artists	bakers – chemists

Table 2: The four classical figures.

In this paper we present a novel approach by applying a generic cognitive characterization of (i) the reasoners and of (ii) the tasks. We then investigate whether differences exist by considering real human data, and how these differences can be explained. For this purpose we consider human syllogistic reasoning, as it has been well studied in the past and provides enough variations within tasks.

Syllogisms originate from Aristotle (Barnes, 1984): Each syllogism consists of a pair of syllogistic premises,

[1]The authors are mentioned in alphabetical order.

according to the four classical moods and figures shown in Table 1 and 2, together with one conclusion about a and c expressed in one of the four moods in Table 1. There are $4^2 \times 4 = 64$ different pairs of premises that can be uniquely specified by the abbreviations of the moods and figures in Table 1 and 2. Consider the following pair of syllogistic premises, abbreviated by AI1 (Aab, Ibc):

All artists are bakers.   Some bakers are chemists.

Given these two premises, which conclusion on the relation between artists and chemists (necessarily) follows? A (syllogistic) conclusion can take one of the following forms:

Aac   Eac   Iac   Oac   Aca   Eca   Ica   Oca

where the abbreviation Aac stands for *All a are c*, Eac stands for *No a are c*, etc. According to classical logic, *no valid conclusion* (NVC) follows, while according to the data in (Khemlani & Johnson-Laird, 2012), 70% of the participants concluded *Some artists are chemists* (Iac) and 16%, which was just above random choice, concluded NVC. Here, we are not only interested in these individual differences but investigate whether the same participants change their underlying assumptions when carrying out the task of solving these 64 pairs of (syllogistic) premises.

The paper is structured as follows: We will first introduce the relevant cognitive principles. After that, we will present the novel computational approach of *clustering by principle patterns* for syllogistic reasoning. In particular, the generic cognitive characterization of (i) reasoners and of (ii) tasks will be developed. Finally, a two-folded evaluation is given: First, the predictive accuracy of *clustering by principle patterns* will be compared to the benchmark models with the help of the CCOBRA framework. Second, a qualitative assessment will give more insight on the relation between reasoners, tasks and principles.

Cognitive Principles

The underlying theory of *clustering by principle patterns*, is the weak completion semantics (WCS) (Hölldobler & Kencana Ramli, 2009), a logic programming approach (Lloyd, 1984). Tasks within WCS are represented formally as logic programs, and reasoning takes place with respect to their supported models that are computed by a semantic fixed point operator interpreted under the three-valued Łukasiewicz (1920) logic. Whenever applicable, explanations are skeptically abduced and programs minimally revised. WCS has already been applied for various reasoning tasks, such as Byrne's (1989) suppression task (Dietz, Hölldobler, & Ragni, 2012), Wason's (1968) selection task (Dietz, Hölldobler, & Ragni, 2013), and human syllogistic reasoning (Dietz, 2017; Costa, Dietz Saldanha, Hölldobler, & Ragni, 2017). Here, we will not introduce the underlying formalization within WCS, but rather give an intuitive understanding of the principles.

Interpretation of conditionals and quantifiers As widely accepted, humans understand natural language statements, such as *All bakers are artists* defeasibly (Stenning & van Lambalgen, 2005, 2008). Accordingly, for any syllogistic premise, we keep the structure of a conditional (**conditionals**) but include a license for inference (**licenses**), by means of an abnormality predicate, as follows:

All bakers, that are not abnormal, are artists.
By default, no baker is abnormal.

Additionally, humans do not quantify over things that do not exist (Grice, 1975), which is called the existential import (**import**). Thus, if we know that all *All bakers are artists*, we assume that *Some bakers exist.* In the sequel, for all universally quantified premises (A and E), (conditionals), (licenses) and (import) apply.

According to Grice (1975), ideally, humans try to be as informative and clear as possible. Thus, *Some bakers are artists* implies that *not all bakers are artists*, because otherwise it would have been stated. For instance, for *Some bakers are artists*, we assume that *For some bakers, it is unknown whether they are artists* (**unknownGen**). Even though in classical logic *Some bakers are artists* is equivalent to *Some artists are bakers*, not all humans make that assumption, and therefore we explicitly introduce the (**converse**) principle when assuming this equivalence. All introduced principles so far belong to the **basic** principles.

Additional Inferences and Generalizations Some humans generalize over premises with the existentially quantified moods, I (*Some ... are ...*) and O (*Some ... are not ...*). They seem to understand them as their corresponding (limited) universally quantified moods, A and E, respectively with respect to entities that are not directly introduced by the premise in consideration (**generalization**). Even though it seems unlikely that humans reason directly by contraposition while solving a reasoning task, humans might come to the contrapositive conclusion by searching for counter examples (cf. (Rips, 1994; O'Brien, D. S. Braine, & Yang, 1994)) (**contraposition**). On the other hand, if participants cannot straightforwardly derive any conclusion, instead of deriving NVC, they search for explanations (**searchAlt**) (Lipton, 2003). Under WCS, (searchAlt) is modeled by means of skeptical abduction.

An overview of the predictions of these principles for all 64 pairs of syllogistic premises within WCS can be found in (Dietz Saldanha & Mörbitz, 2020, Table 5).

Task	Part A	Explaining Principles	Part B	Explaining Principles	Part C	Explaining Principles
AI1	Iac	abd_1, atmo_1	Iac	abd_1, atmo_1	Ica	abd_2, atmo_2
AI3	Ica	abd_2, atmo_2	NVC	basic, contrap, general	NVC	basic, contrap, general
AE3	Eca	contrap, atmo_2	Eca	contrap, atmo_2	Eca	contrap/atmo_2
EI3	Oca	contrap, atmo_2	Oca	contrap, atmo_2	NVC	basic, abd
IA2	Ica	abd_2, atmo_2	NVC	basic, contrap	Ica	abd_2, atmo_2
II1	Iac	atmo_1	NVC	basic, contrap, abd	NVC	basic, contrap, abd
	Atmosphere (6 out of 6)		Contraposition (5 out of 6)		various	

Table 3: Three participants' (A, B and C) responses for six tasks and the principles that explain these responses.

Heuristic strategies The following two principles describe heuristic strategies applicable to syllogistic reasoning. According to the **atmosphere** bias (Woodworth & Sells, 1935), humans might be affected by the moods of the premises, in the sense that universal (affirmative, resp.) conclusions are excluded when one of the premises is existential (negative, resp). In the case of identical moods, the conclusion must have this mood as well. Riesterer, Brand, Dames, and Ragni (2019) observed that most cognitive theories do not predict NVC. Therefore, they develop own NVC heuristics, for which the one with the highest predictive accuracy is specified as follows: If none of the premises contains the affirmative universal quantifier *for all*, NVC is predicted (**nvc heuristic**).

Three Participants and their Explaining Principles As illustration for the introduced cognitive principles applied to syllogistic reasoning consider Table 3: Column 1 shows the given task (pair of premises), and column 2 to 4 show the cognitive principles that correspond to the responses given by participant A, B, and C, respectively.[2] basic, contrap, general and abd are abbreviations for the principles basic, contraposition, generalization and abduction, respectively. Note that basic is assumed by contrap, general and abd as well.

In case a principle predicts two responses, $_1$ refers to the corresponding ac response whereas $_2$ to the ca response. The last row shows which of the principles explained most of the participants' responses for the six tasks: A's and B's responses are best explained by the atmosphere and contraposition principles, respectively, whereas for C's responses no preferred principles seem to apply.

Clustering by Principle Patterns

We are interested in finding the clusters which best fit the response patterns of participants within a certain population. The idea originates from *Clustering by Principles* proposed in (Dietz Saldanha & Schambach, 2019). This approach assumes that a population of participants can be characterized by a set of clusters, where each of them is characterized by a set of principles. However, here we intend to compute clusters that should not only be characterized by a restricted set of underlying principles, but rather, their characterization depends on the given task.

For this purpose, we first applied the machine learning k-means++ clustering algorithm (MacQueen et al., 1967; Arthur & Vassilvitskii, 2006) to the task, where the distance metric was adapted to fit the participants' data. A data point is a response pattern given by one participant, i.e. it consists of all 64 responses, where each belongs to one of the 64 tasks (pairs of syllogistic premises). For instance, $[\mathsf{Aac}, \mathsf{NVC}, \mathsf{Aca}, \ldots, \mathsf{Aca}]$ of length 64 is a valid data point. The *distance* between two datapoints is calculated using the hamming distance between their answers, which is computed by comparing the number of different responses, and computing the sum of the number of different answers. Consider two data points of length 64, where one only consists of 64 NVC responses and the other one only consists of 64 Aac responses: The computed distance is 64, as they have 64 different responses. The distance of a data point to itself is 0, as all answers are the same.

Using this metric, the algorithm computes k different data points, i.e. clusters, with a corresponding cluster center. This cluster center does not necessarily need to correspond to the response pattern but it has the smallest distance to all of the data points in its cluster in comparison to all the other k-1 center points.

These cluster centers represent the k-clusters and are then compared to the responses predicted by the cognitive principles. Hence, for each cluster center, we specify a set of 64 pairs, where one pair consists of the task (pair of syllogistic premises) and a cognitive principle that predicted the same answer, i.e. the answer that explains the response of each task of the center.

As illustration consider the principles that explain participant's A responses in Table 4. Let us assume that the responses given by participant A is a cluster center. For OA1, IA4 and AI2, the participant's responses match (among others) the basic principle's prediction, whereas

[2]The data containing these responses is in Ragni2016.csv in `https://github.com/CognitiveComputationLab/ccobra`

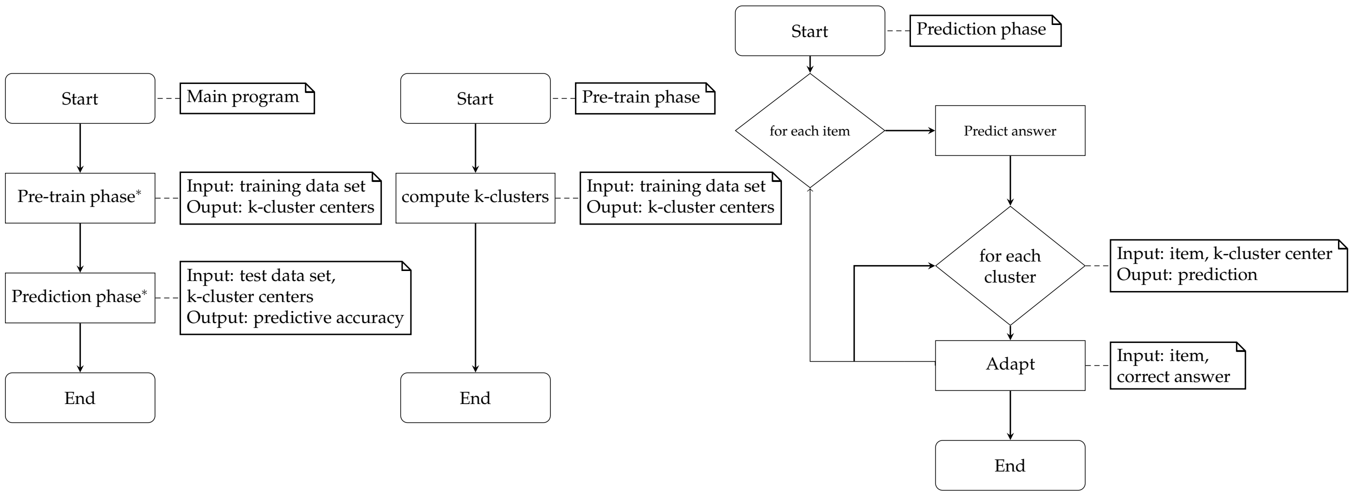

Figure 1: Flowchart of program (left), pre-train (middle) and prediction (right).

Task	Part A	Principles Explaining A's Response
OA1	Oac	basic, contrap, general, abd, atmo_1
IA4	Ica	basic_2, contrap_2, general_2, abd_2, atmo_2
AI2	Iac	basic_1, contrap_1, general_1, abd_1, atmo_1
EA3	Oac	none
AA1	Ica	none
EI4	NVC	nvc heuristic
EO3	NVC	basic, contrap, general, abd, nvc heuristic
OO1	NVC	basic, contrap, general, abd, nvc heuristic
AA4	Iac	none
OA4	Oac	atmo_1
IO4	Oca	atmo_2
AE3	Eca	atmo_2

Table 4: Extract of response pattern for participant A.

for OA4, IO4, and AE3 it matches the predictions made by the atmosphere bias. Yet, the participants' answer for EI4, EO3 and OO1 corresponds to the nvc heuristics. A cluster center that corresponds to participants' A pattern could be a set containing these pairs:

$$\{\ldots,(\text{OA1},\text{basic}),\ (\text{IA4},\text{basic}_2),\ (\text{AI2},\text{basic}_2),\\ (\text{EI4},\text{nvc heuristic}),\ (\text{EO3},\text{nvc heuristic}),(\text{OO1},\text{nvc heuristic}),\\ (\text{OA4},\text{atmo}_1),\ (\text{IO4},\text{atmo}_2),\ (\text{AE3},\text{atmo}_2),\ldots\}$$

Note that the same answer can be predicted by different principles (e.g. EO3 and OO1 are also predicted by contrap, general, abd and nvc heuristic), and thus different pairs can lead to the same predictions.

Implementation in CCOBRA

The evaluation environment for the predictive accuracy, called the CCOBRA (Cognitive COmputation for Behavioral Reasoning Analysis) Modeling Framework, in which a given cognitive model can adapt the optimal prediction strategy through the pre-train phase, by dynamically changing its predictions depending on the participants' past responses. The modeling framework can be found here: `https://github.com/CognitiveComputationLab/ccobra`.

Figure 1 (left) shows an overview of our implementation within the CCOBRA framework, that consists of two main phases: The pre-train phase (middle) and the prediction phase (right). In the pre-train phase, the model is asked to predict the responses from the training data set, and then to adapt its own strategy according to their responses. For this phase, the k-means++ clustering algorithm is applied to the training data set. The prediction using these k-cluster centers is implemented as follows: First, for each k-cluster center that predicts the participants' response correctly during the adapt phase, its corresponding predictive accuracy (score) is increased dynamically. Initially, when no information about the participants' response pattern is known, the predicted answer will be the one corresponding to a k-cluster computed with k-means, where k = 1. This k-cluster center has the minimum distance to all data points, and therefore its response corresponds to the most frequent answers (MFA) given by the training data set. After each prediction, the k-cluster center that matched the last responses of the participants best, is chosen for the next prediction. An elaborate description of the implementation within CCOBRA is provided in (Dietz Saldanha & Schambach, 2019). The implementation of *clustering by principle patterns* is accessible online: `https://github.com/enterJazz/syllogistic-k-means`.

Evaluation

The evaluation was done in two parts: First, the predictive accuracy of the clustering method was compared

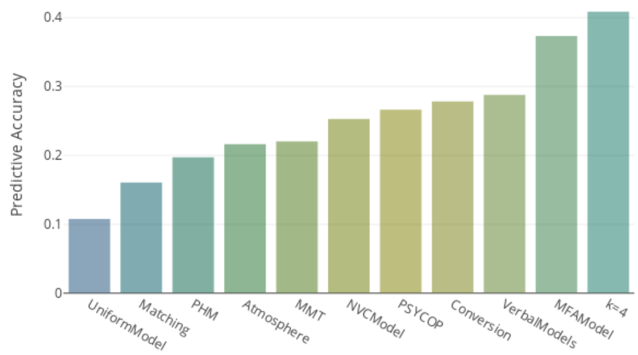

Figure 2: Predictive accuracy of benchmarks and *Clustering by Principle Patterns* (rightmost).

with the performance of the available benchmark models. As training data, Ragni2016.csv provided by CCOBRA was used, which contained the response patterns of 139 participants for all 64 tasks. As test data for the prediction phase, Veser2018.csv from CCOBRA was used, which contained the response patterns of 32 participants. We will also discuss the predictions under *clustering by principle patterns* in combination with the nvc heuristics from (Riesterer et al., 2019).

After that, a qualitative analysis is carried out, by investigating whether a correspondence between individuals, tasks and principles could be observed.

Predictive Accuracy

Figure 2 shows the results of the evaluation of *clustering by principle patterns* assuming 4 clusters, i.e. where k=4, compared to the predictions of the benchmark model provided by CCOBRA. From left to right the figure shows the results of the following models: uniform (11%, when the responses are chosen randomly), matching bias (16%, matching) (Wetherick & Gilhooly, 1995), probability heuristics model (20%, PHM) (Chater & Oaksford, 1999), atmosphere bias (22%, atmo) (Woodworth & Sells, 1935), mental model theory (22%, MMT) (Johnson-Laird, 1983), NVC (25%, when the chosen responses are always *no valid conclusion*), (logic-based) PSYCOP model (27%, PSYCOP) (Rips, 1994), illicit conversion heuristics (28%, Conversion) (Chapman & Chapman, 1959), verbals models theory (29%, VerbalModels) (Polk & Newell, 1995), MFA (37%, when the most frequent answer learned from the training data set is chosen) and *clustering by principles patterns* (41%, k=4). Note that Figure 2 has been automatically generated by the CCOBRA framework.

Clustering by principle patterns achieved a predictive accuracy of just above 40%, which is the highest score, compared to the performance of the other benchmark models. Note that when applying cross validation on the training set, the predictive accuracy is just above 50%, which is about the same result as other machine learning approaches have achieved so far.[3]

NVC heuristics

According to Riesterer et al. (2019), most cognitive theories do not account for NVC conclusions. Therefore, we have tested all the NVC heuristics presented in (Riesterer et al., 2019) in combination with *clustering by principle patterns* but no overall improvement was observed, the predictive accuracy was even worse.

From a qualitative point of view the following can be observed: The predictions under WCS guided by the cognitive principles covered 33 out of the 36 NVC conclusions made by the best performing heuristic in (Riesterer et al., 2019). This seems to indicate the these predictions accounted well for NVC conclusions in syllogistic reasoning.

Observations on Individuals, Tasks & Principles

Are there variations among (i) individuals and (ii) tasks that can be characterized by clusters explained by cognitive principles? Can we observe a certain tendency of occurring principles within clusters, and if so, how can these clusters be characterized? Let us first consider Table 5, where for six tasks, columns 2 to 5 show the predictions of the four learned clusters. The gray percentages in brackets (in the first row) show how many of the participants in the training set corresponded to that particular cluster: Almost half of the participants (45%) have been assigned to cluster 1, whereas only 13% have been assigned to cluster 4. As the predictions of each cluster show, none of them can be characterized by a particular principle, but rather, the corresponding principles depend on the given task. The predictions that correspond to the highest achieved percentage are highlighted in gray. Column six shows which principles explain these predictions. For instance, for task II1, $\mathsf{general}_1$ explains the response Iac.

The last column of Table 5 shows which of the benchmark models predicted the same responses (the ones highlighted in gray). Interestingly, atmo, MMT, VerbalModels and Matching predicted all the highlighted responses, without distinguishing among participants. Our original claim that the applicable principles differ throughout all clusters, is confirmed. We conjecture that the following principles are preferably applied within particular tasks: (i) Contraposition seems preferably applied when both premises are universally quantified (AE3 vs. EI3) and (ii) Abduction seems preferably applied on syllogistic premise pairs of Figure 1 and 2 (AI1 and IA2 vs AI3). However, further investigations are necessary to confirm these claims.

[3] https://www.cc.uni-freiburg.de/staff/files/2019-04-dresden

Task	1 (45%)	2 (19%)	3 (23%)	4 (13%)	Explaining Principle	Theories with same prediction
AI1	Iac	Iac	Iac	Iac	abd_1 (100%)	PHM, MMT, VerbalModels, Conversion, Matching, atmo
AI3	NVC	Iac	Ica	Ica	abd_2 (36%)	PHM, MMT, VerbalModels, Conversion, Matching, atmo
AE3	Eac	Eca	Eca	Eac	contrap_1 (58%)	MMT, PSYCOP, VerbalModels, Conversion, Matching, atmo
EI3	NVC	Oca	Eac	Oca	$\text{contrap}_{1,2}$ (32%)	MMT, PSYCOP, VerbalModels, atmo
IA2	Ica	Ica	Ica	Ica	abd_2 (100%)	PHM, MMT, VerbalModels, Matching, atmo
II1	NVC	Iac	Iac	Iac	general_1 (55%)	PHM, MMT, VerbalModels, Matching, atmo

Table 5: Four Clusters and their predictions for six tasks. Column 6 shows the principle that explains the (gray) highlighted response in the respective row and the last column shows the theories with the same predictions.

Similar to Johnson-Laird and Khemlani (2016)'s three identified groups of reasoners, deliberative, intermediate, and intuitive, we can characterize the developed clusters here according to the frequency and types of principles they apply: Over 80% of cluster 1 responses can be explained by abduction and contraposition, whereas for cluster 2 responses, over 80% can be explained by the atmosphere bias. Accordingly, it seems reasonable to classify the participants in cluster 1 as deliberative reasoners, whereas the participants in cluster 2 might best fit the intuitive reasoners. The responses given in cluster 3 and 4 seem not to have a strong tendency for any set of principles. This is consistent with the observation made in (Dietz Saldanha & Mörbitz, 2020): For some participants' reasoning patterns, a straightforward characterization by means of the principles (identified so far) seems to be difficult.

Conclusions

We investigated which responses of the cognitive principles within human syllogistic reasoning are applicable given a certain task (pair of syllogistic premises). With *clustering by principle patterns* we have developed a generic cognitive characterization which shows that cognitive theories do not only need to account for (i) differences among reasoners, but also for (ii) differences among tasks. The best fitting clusters were computed by applying the k-means++ clustering algorithm on the training data set and explained by the principles that predicted the same responses. Each of these clusters represent a response pattern for all 64 tasks and outperformed the benchmark models. More importantly, as these clusters are characterized by underlying explainable and cognitively plausible principles, we can observe that principles are preferably applicable not only with respect to a certain cluster, but with respect to a given task. Finally, by considering the amount of principles which fitted the responses of each cluster, we were able to characterize cluster 1 and 2. The predictions given by the benchmark theories in the last column of Table 5 do not account for neither (i) or (ii).

Summing up, *clustering by principle patterns* is novel and contributes in various aspects: The approach integrates the clusters computed by a machine learning algorithm with underlying cognitive principles, that have their roots in the literature of cognitive science and philosophy of language formalized within WCS. These principles provide cognitively plausible characterizations of the patterns that cover the population of reasoners. We are not aware of other approaches that addressed these differences.

For future directions, the conjectures need to be verified. Possibly new experiment might give explanations when principles, such as abduction or contraposition, are more likely to apply in certain circumstances than in others. It would further help to consider a generic cognitive characterization for other reasoning episodes such as conditional or counterfactual reasoning. These investigations might reveal new types of cognitive principles.

Acknowledgments

Many thanks to Nicolas Riesterer for the helpful and quick replies on our questions concerning CCOBRA. The authors are grateful to the anonymous reviewers for their feedback.

References

Arthur, D., & Vassilvitskii, S. (2006). *k-means++: The advantages of careful seeding* (Tech. Rep.). Stanford.

Barnes, J. (Ed.). (1984). *The complete works of aristotle (vol. 1)*. Princeton, NJ: Princeton University Press.

Breu, C., Ind, A., Mertesdorf, J., & Ragni, M. (2019). The weak completion semantics can model inferences of individual human reasoners. In F. Calimeri, N. Leone, & M. Manna (Eds.), *Logics in artificial intelligence - 16th european conference, JELIA 2019, rende, italy, may 7-11, 2019, proceedings* (Vol. 11468, pp. 498–508). Springer.

Byrne, R. M. (1989). Suppressing valid inferences with conditionals. *Cognition*, *31*, 61–83.

Chapman, L. J., & Chapman, J. P. (1959). Atmosphere effect re-examined. *Journal of Experimental Psychology*, *58*(3), 220-6.

Chater, N., & Oaksford, M. (1999). The probability heuristics model of syllogistic reasoning. *Cognitive Psychology*, *38*(2), 191 - 258.

Costa, A., Dietz Saldanha, E., Hölldobler, S., & Ragni, M. (2017). A computational logic approach to human syllogistic reasoning. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of 39th conference of cognitive science society* (pp. 883–888). Austin, TX: Cognitive Science Society.

Dietz, E.-A. (2017). A computational logic approach to the belief bias in human syllogistic reasoning. In *Proceedings of 10th international and interdisciplinary conference on modeling and using context* (Vol. 10257, p. 691-707). Springer.

Dietz, E.-A., Hölldobler, S., & Ragni, M. (2012). A computational logic approach to the suppression task. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of 34th conference of cognitive science society* (pp. 1500–1505). Austin, TX: Cognitive Science Society.

Dietz, E.-A., Hölldobler, S., & Ragni, M. (2013). A computational logic approach to the abstract and the social case of the selection task. In *Proceedings of the 11th international symposium on logical formalizations of commonsense reasoning, (Commonsense 2013).* Aeya Nappa, Cyprus.

Dietz Saldanha, E., & Schambach, R. (2019). Human syllogistic reasoning: Towards predicting individuals' reasoning behavior based on cognitive principles. In C. Beierle, M. Ragni, F. Stolzenburg, & M. Thimm (Eds.), *Proceedings of the 8th workshop on dynamics of knowledge and belief and the 7th workshop KI & kognition* (Vol. 2445, pp. 2–13). CEUR-WS.org.

Dietz Saldanha, E., & Mörbitz, R. (2020). Cognitive Principles and Individual Differences in Human Syllogistic Reasoning. In M. Urbański, T. Skura, & P. Łupkowski (Eds.), *Reasoning: Games, cognition, logic, college publications* (pp. 59 – 80). College Publications, London.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Vol. 3: Speech acts* (p. 41-58). New York: Academic Press. (Reprinted as ch.2 of Grice 1989, pages 22–40)

Hölldobler, S., & Kencana Ramli, C. D. P. (2009). Logic programs under three-valued Łukasiewicz semantics. In P. M. Hill & D. S. Warren (Eds.), *Proceedings of 25th international conference on logic programming* (Vol. 5649, pp. 464–478). Springer Berlin Heidelberg.

Johnson-Laird, P. N., & Khemlani, S. S. (2016). How people differ in syllogistic reasoning. In G. D. M. D. Papafragou A. & J. Trueswell (Eds.), *Proceedings of 38th conference of cognitive science society* (pp. 2165–2170). Austin, TX: Cognitive Science Society.

Johnson-Laird, P. N. (1983). *Mental models: towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.

Khemlani, S., & Johnson-Laird, P. N. (2012). Theories of the syllogism: A meta-analysis. *Psychological Bulletin*, 427-457.

Lipton, P. (2003). *Inference to the best explanation*. Routledge.

Lloyd, J. W. (1984). *Foundations of logic programming*. New York, NY, USA: Springer-Verlag New York, Inc.

Łukasiewicz, J. (1920). O logice trójwartościowej. *Ruch Filozoficzny*, *5*, 169–171. (English translation: On three-valued logic. In: Łukasiewicz J. and Borkowski L. (ed.). (1990). *Selected Works*, Amsterdam: North Holland, pages 87–88.)

MacQueen, J., et al. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth berkeley symposium on mathematical statistics and probability* (Vol. 1, pp. 281–297).

O'Brien, D., D. S. Braine, M., & Yang, Y. (1994, 11). Propositional reasoning by mental models? simple to refute in principle and in practice. , *101*, 711-24.

Polk, T. A., & Newell, A. (1995). Deduction as verbal reasoning. *Psychological Review*, *102*(3), 533-566.

Ragni, M., Kola, I., & Johnson-Laird, P. (2017). The wason selection task: A meta-analysis. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of 39th conference of cognitive science society* (pp. 980–985). Austin, TX: Cognitive Science Society.

Riesterer, N., Brand, D., Dames, H., & Ragni, M. (2019). Modeling human syllogistic reasoning: The role of "no valid conclusion". In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of 41th conference of cognitive science society* (pp. 953–959). cognitivesciencesociety.org.

Rips, L. J. (1994). *The psychology of proof: Deductive reasoning in human thinking*. Cambridge, MA: MIT Press.

Stenning, K., & van Lambalgen, M. (2005). Semantic interpretation as computation in nonmonotonic logic: The real meaning of the suppression task. *Cognitive Science*, *6*(29), 916–960.

Stenning, K., & van Lambalgen, M. (2008). *Human reasoning and cognitive science*. Cambridge, MA: MIT Press.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*(3), 273–281.

Wetherick, N. E., & Gilhooly, K. J. (1995). 'atmosphere', matching, and logic in syllogistic reasoning. *Current Psychology*, *14*(3), 169–178.

Woodworth, R. S., & Sells, S. B. (1935). An atmosphere effect in formal syllogistic reasoning. *Journal of Experimental Psychology*, *18*(4), 451.