

Arbitrarily Applicable Relational Responding in NARS

Robert Johansson^{1,2}

¹ Department of Computer and Information Science, Linköping University, Sweden

² Department of Psychology, Stockholm University, Sweden

`robert.johansson@liu.se`

Abstract. In a previous publication, we argued why the behavioral psychology theory *Relational Frame Theory (RFT)* might be interesting for AGI researchers. This paper explores details of RFT in NARS. More specifically, we investigate different response patterns, such as equivalence, opposition, and comparison in NARS. An additional core feature of RFT is *transformation of stimulus function* which explains how arbitrary symbols can acquire various functions depending on the relational networks in operation. We demonstrate how NARS handles this process. Finally, future applied and basic research opportunities are discussed.

Keywords: Relational Frame Theory · NARS · Intelligence · Self.

1 Introduction

In a recent paper [3], we argued why the theory *Relational Frame Theory (RFT)* [1] might be interesting for researchers in the field of Artificial General Intelligence. RFT provides a unified account of language and cognition from the perspective of behavior psychology. The key thesis of RFT is that intelligence and cognitive functions are all instances of *arbitrarily applicable relational responding (AARR)*. AARR is defined as higher-order response patterns that fulfill a specific set of criteria. Specific instances of AARR such as equivalence, opposition, and comparisons are referred to as different types of *relational frames*. RFT suggests that key properties of relational frames are *mutual entailment*, *combinatorial entailment*, and *transformation of stimulus function*. Furthermore, relational frames are *contextually controlled* and developed through *multiple-exemplar training*. These concepts are explained more thoroughly in the original publication [3]. The aim of this paper is to explore these RFT concepts in NARS.

2 Stimulus equivalence and derived stimulus relations

One of the foundations of RFT is the link to the *stimulus equivalence* phenomenon [4]. In essence, it is one of the most basic forms of AARR, as will be explained later in this paper. Stimulus equivalence can be observed in verbally able humans, but seems to be very difficult, if not impossible, in non-humans.

A simple example follows. Let's say that someone learns that “*A is the same as B, and B is the same as C*”, with A, B, and C being arbitrary symbols or events. We can label this as that the person learns the AB and AC relations. Then, the person will without further training be able to derive the BA, CB, AC and CA relations. Hence, the stimuli have become *equivalent*. In the stimulus equivalence literature, deriving BA and CB are examples of *derived symmetry* or just *symmetry*, while deriving AC and CA are examples of *transitivity*. The stimulus equivalence phenomenon is a very important finding, as it involves basic examples of *derived stimulus relations*.

In NARS, with its focus on derived relations, the capacity for forming equivalence relations is a built-in feature. The system can do that without previous training, as showed below.

```
<A1 <-> B1>.
<A1 <-> C1>.
<C1 <-> B1>?
Answer: <B1 <-> C1>. %1.00;0.81%
```

To clarify, in NARS, this kind of relating is not learned from experience. Such learned relations are explored below.

3 AARR and relational frames

As described above, the form of relating involved in stimulus equivalence is an example of AARR. Importantly though, AARR involves many other forms of relating beyond equivalence. For example, if a verbally able human learns that “*A is larger than B, and B is larger than C*”, (AB and AC), the person will be able to derive that “*B is smaller than A*” (BA), “*C is smaller than B*” (CB), “*A is larger than C*” (AC), and “*C is smaller than A*” (CA). In RFT terms, it is said that the BA and CB relations are *mutually entailed*, and the AC and CA relations are *combinatorially entailed*. These terms are more general ways of talking about symmetrical and transitive relations, respectively.

Different instances of AARR are named as different *relational frames*. The most common relational frames mentioned in the literature are *coordination* (sameness), *opposition*, *distinction*, *comparison* (more/less than), *hierarchy*, *temporal*, *spatial*, and *deictic* (“*I am here now*” vs “*You were there then*”).

All these relational frames (and others) have the properties of mutual and combinatorial entailment, as described above. Other important features of relational frames (thoroughly described later in this paper) are that they develop through *multiple-exemplar training*, they have the property of *transformation of stimulus function*, and that they are under *contextual control*.

4 Relational frames in NARS

Complex relations beyond equivalence can be learned through experience in NARS. In the following examples, NARS is given experiences of opposition and comparison relations.

```

<(*,A1,B1) --> opposite>. :|:
<(*,B1,A1) --> opposite>. :|:
<<(*,$1,$2) --> opposite> <=> <(*,$2,$1) --> opposite>>?
Answer: <<(*,$1,$2) --> opposite> <=>
        <(*,$2,$1) --> opposite>>. %1.00;0.22%

<(*,A2,B2) --> more>. :|:
<(*,B2,A2) --> less>. :|:
<<(*,$1,$2) --> more> <=> <(*,$2,$1) --> less>>?
Answer: Answer <<(*,$1,$2) --> more> <=>
        <(*,$2,$1) --> less>>. %1.00;0.22%

<(*,A3,B3) --> opposite>. :|:
<(*,B3,A3) --> opposite>?
Answer: <(*,B3,A3) --> opposite>. :-443: %1.00;0.27%

<(*,A4,B4) --> less>. :|:
<(*,B4,A4) --> more>?
Answer: Answer <(*,B4,A4) --> more>. %1.00;0.12%

```

As seen in the example, despite having only been given a single experience of the bidirectionality of relations, NARS generates a hypothesis of a more general pattern. With more examples, NARS will be more confident. The fact that NARS can learn a relational response using different examples illustrates *multiple-exemplar training* in NARS.

Given that NARS has learned the general pattern (through multiple-exemplar training, or direct learning of the abstract pattern), the system can then provide a mutually entailed response, such as `(*,B3,A3) --> opposite>`. Mutually entailed responses for different relations learned from experience works well in the current version of NARS.

The next level of complexity in derived stimulus relations are combinatorially entailed relations. For example, in the case of a comparative relation, the derived AC relation is the same as the AB and the BC relations. Below is shown how NARS can learn this from experience.

```

<(*,A1,B1) --> more>. :|:
<(*,B1,C1) --> more>. :|:
<(*,A1,C1) --> more>. :|:

<(&|,<(*,$1,#2) --> more>,<(*,#2,$3) --> more>) =|>
  <(*,$1,$3) --> more>>? :|:
Answer: <(&|,<(*,#1,$2) --> more>,<(*,$3,#1) --> more>) =|>
        <(*,$3,$2) --> more>>. :|: %1.00;0.34%

```

NARS can also start out with knowledge of transitivity, like in the example below.

```

<&&&<(*,$1,#2) --> more>,<(*,#2,$3) --> more> ==>
  <(*,$1,$3) --> more>>.
<(*,A1,B1) --> more>.
<(*,B1,C1) --> more>.
<(*,A1,C1) --> more>?
Answer: <(*,A1,C1) --> more>. %1.00;0.05%

```

Importantly, in the current NARS version (3.0.2) there is a problem with transitive relations beyond the examples above. For example, if the latter example would be followed by $\langle(*,A2,B2) \text{ --> more}\rangle$ and $\langle(*,B2,C2) \text{ --> more}\rangle$, the current NARS wouldn't typically derive $\langle(*,A2,C2) \text{ --> more}\rangle$, despite its experiences. This is due to a set of design decisions in the current NARS control mechanism.

In summary, NARS seems at its current stage to be able to form mutually entailed relations through multiple-exemplar training. It is also clear in the NARS logic how combinatorially entailed relations would work, even if it at the current stage doesn't work completely as expected.

5 Multiple stimulus relations

One cornerstone of RFT is how multiple stimulus relations combine into relational networks. In the study by Steele and Hayes [5], the authors showed this behavior experimentally for the first time. See the below Figure for an illustration of the relational network trained. This is also a very important paper for theoretical reasons. Up until this paper, the stimulus equivalence phenomenon had been explained by equivalence classes being formed for the relevant stimuli [4]. However, the fact that AARR can be both non-symmetric and non-transitive, has lead to the RFT field questioning the value of a class-based analysis.

The Steele and Hayes network can be encoded in NARS as follows, with the questions being answered one at a time.

```

<<(*,$1,$2) --> opposite> <=> <(*,$2,$1) --> opposite>>.
<A1 <-> B1>.
<A1 <-> C1>.
<(*,A1,B2) --> opposite>.
<(*,A1,C2) --> opposite>.
<(*,C2,D1) --> opposite>.
<C2 <-> D2>.

<(*,D1,B2) --> opposite>?
<D1 <-> B1>?
<(*,A1,D2) --> opposite>?
<A1 <-> D1>?

Answer: <(*,D1,B2) --> opposite>. %1.00;0.31%
Answer: <B1 <-> D1>. %1.00;0.37%

```

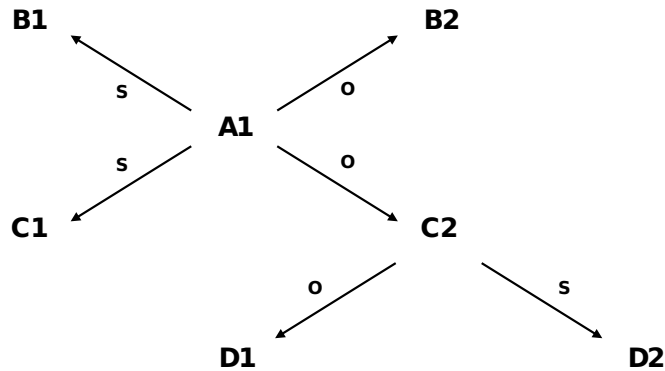


Fig. 1. The relations trained in the study by Steele and Hayes [5]. Subjects were trained in SAME (S) and OPPOSITE (O) relations, and tested on various combined derived relations.

Answer: $\langle (*, A1, D2) \rightarrow \text{opposite} \rangle$. %1.00;0.73%
 Answer: $\langle A1 \leftrightarrow D1 \rangle$. %1.00;0.42%

In the example above, the combinatorially entailed SAME response $D1-B2$ is working without NARS having been taught such a response. This is due to a fact that the control mechanism of NARS in some situations “defaults” to a SAME response. Importantly though, the control mechanism of NARS would need more work to be able to combine relations reliably. However, there is nothing in the NARS logic itself that prevents this.

This example also highlights the *contextual control* cornerstone of RFT. One importance aspect of contextual control is that the responses are controlled by contextual information. This can be seen in the example above, in that the system’s responses are a function of the relational experiences provided to the system. Another important aspect of contextual control is how the context can come to control functions of stimuli. This is described in the next section.

6 Transformation of stimulus function

At the heart of RFT is the *transformation of stimulus function*. The term *stimulus function* is used very broadly, and covers all potential roles that stimuli can have in various contexts. This can be for example perceptual functions, like in the case of “*Bananas are yellow*” or “*Apples taste sweet*”, but also appetative (reinforcing) or aversive (punishing) functions. For example, electric shocks have aversive functions for human beings, due to their genetic setup. *Transformation of stimulus function* refers to the fact that functions for one stimulus can be transformed based on its relations to other stimuli. A simple example in NARS follows.

```

<A1 <-> B1>.
<A1 <-> C1>.
<B1 --> [red]>.
<C1 --> [?1]>?
Answer: <C1 --> [red]>. %1.00;0.73%

```

In the above example, a simple perceptual function is transferred from B1 to A1 and C1. Importantly, the relation between stimuli doesn't need to be equivalence, as illustrated by the example below.

```

<(*,[cold],[hot]) --> opposite>.
<(*,A1,B2) --> opposite>.
<(*,A1,C2) --> opposite>.
<B2 --> [hot]>.

<C2 --> [?1]>?
<A1 --> [?1]>?
Answer: <C2 --> [hot]>. %1.00;0.31%
Answer: <A1 --> [cold]>. %1.00;0.45%

```

The above example uses the property copula from NARS to describe stimulus functions. While this works for simple examples, it can't capture a statement such as “*A is of opposite temperature to B*”. If one views a stimulus function as a *relation* rather than a *property*, an alternative NARS encoding is natural. For example, <A --> [hot] (hot is a property of A) can be described as <(*,A,[hot]) --> temperature (A and hot is in a relation called temperature). Given this, we can construct more complex examples.

```

<(*,[cold],[hot]) --> opposite>.
<(*,A,[cold]) --> temperature>.
<(*,A,[blue]) --> color>.
<(*,(/,temperature,A,_),(/,temperature,B,_)) --> opposite>.
<(/,color,A,_) <-> (/,color,B,_)>.

<(*,B,[?1]) --> temperature>?
<(*,B,[?1]) --> color>?
Answer: <(*,B,[blue]) --> color>. %1.00;0.81%
Answer: <(*,B,[hot]) --> temperature>. %1.00;0.45%

```

7 Other relational frames in NARS

Three additional frames can be said to be built-in into NARS, *hierarchy*, *distinction* and *temporal* frames. *Hierarchy* can simply be represented by the inheritance copula in NARS, as illustrated below.

```

<fruit --> food>.

```

```

<apple --> fruit>.
<food --> [eatable]>.
<apple --> [juicy]>.
<apple --> [eatable]>?
<food --> [juicy]>?
Answer: <apple --> [eatable]>. %1.00;0.73%
Answer: <food --> [juicy]>. %1.00;0.29%

```

The system's response that apples probably are eatable, can be said to be a combinatorially entailed hierarchical response that involves the transfer of the eatable function from food to apples. NARS does also see it as a slight possibility in that food is juicy. Given the experiences above, we can't know for sure.

Regarding the frame of *distinction*, a simple example follows.

```

<A1 <-> B1>.
<A1 <-> B2>. %0%
<B1 --> [dangerous]>.
<A1 --> [dangerous]>?
<B2 --> [dangerous]>? // No answer
Answer: <A1 --> [dangerous]>. %1.00;0.81%

```

In the example above, <A1 <-> B2>. %0% it is shown how it is possible to describe for NARS that two stimuli A1 and B2 are distinct from each other. As expected, NARS can't derive whether B2 is dangerous or not, given that A1 is. This is in line with how RFT describes that transformation of stimulus function works for the frame of distinction.

Temporal frames are also part of NARS from the start, as illustrated with the following example.

```

<tuesday =/> wednesday>.
<wednesday =/> saturday>.
<saturday =\> tuesday>?
Answer: <saturday =\> tuesday>. %1.00;0.45%

```

Furthermore, experiences presented to NARS over time, can be used in temporal reasoning.

```

<{cloud} --> [dark]>. :|:
1000
<{sky} --> [raining]>. :|:
1000
<{SELF} --> [wet]>. :|:
<<{cloud} --> [dark]> =/> <{SELF} --> [wet]>>?
Answer: <<{cloud} --> [dark]> =/>
      <{SELF} --> [wet]>>. %1.00;0.31%

```

Here, NARS makes a combinatorially entailed temporal response, deriving that the system itself may be wet if it observes a cloud that is dark. Regarding transformation of stimulus functions through temporal relations, almost no research exist in this domain.

Regarding *spatial frames*, an example could be that the system learns that “*A is to the left of B, and B is to the left of C*”. A combinatorially entailed spatial response could for example be to look to the right from A, if asked to look at C. In NARS, there has been a recent addition of sensorimotor capabilities [6]. For vision for example, NARS has operations to $\uparrow shift$ and $\uparrow zoom$. This seems like a promising foundation for studying the application of spatial knowledge (frames) as goal-driven action.

Finally, *deictic frames* involves relations in terms of the perspective of the system itself, such as Me/Others, in combination with spatial and temporal relations. For example, a deictic response could be if the system answered “*Green*” given the following information, “*I have a green brick, and you have a red brick. If I was you, and you were me, what brick would you then have?*”. A more complex statement, highlighting the combinatorial nature of the deictic framing, could be “*Yesterday you were sitting there on the black chair, today you are sitting here on the blue chair. If here was there and there was here and if now was then and then was now. Where would be you sitting now?*”.

The recent work on the *SELF* in NARS [7] seems very compatible with deictic framing. NARS learns about itself using operations. In the same way, NARS can also learn about others. This is illustrated below.

```
<(*,{SELF},{green_brick}) --> ^pick>!
<(*,{Michael},{red_brick}) --> ^pick>.
<?1 --> (/ , ^pick, _ , {red_brick})>? // Who picks red brick?
<?1 --> (/ , ^pick, _ , {green_brick})>? // Who picks green brick?

Answer: <{Michael} --> (/ , ^pick, _ , {red_brick})> %1.00;0.90%
Answer: <{SELF} --> (/ , ^pick, _ , {green_brick})> %1.00;0.90%
```

8 Future work

The examples provided in this text have used NARS as a Question-Answering system. With the introduction of procedural operations at the higher layers of NARS, the system can perform goal-achieving actions. At the core, RFT is about viewing relating as a purposeful act [1], rather than that relating takes place “somewhere else” and is then “carried out” by operations. Hence, an important area for future work is to study relational frames as procedural operations.

There are also many future research opportunities regarding transformation of stimulus function in NARS. For example, assume that NARS has been taught to clap and wave, in the presence of stimuli B1 and B2, respectively.

```
<(&/ , <{sample} --> B1> , (^clap, {SELF})) =/>
<{SELF} --> [satisfied]>>.
```



```

<(&/,<{sample} --> B2>,(^wave,{SELF})) =/>
  <{SELF} --> [satisfied]>>.

<{sample} --> B1>. :|:
<{SELF} --> [satisfied]>!
EXE: ^clap([{SELF}])=null

```

In the example above, `<{sample} --> B1>` functions as a *discriminative stimulus*, while `<{SELF} --> [satisfied]>` functions as a *reinforcer* (in the context of the system having a goal of `<{SELF} --> [satisfied]>!`). In an early study by Hayes and co-authors [2], similar behaviors were trained. Then, a relational network was formed where A1 and C1 were taught to be similar to B1. The authors then tested if (and confirmed that) the *discriminative function* of B1 could be transferred to A1 and C1. Similarly, the study showed that *reinforcing functions* also could be transferred. This study seems like a good test when exploring transformation of stimulus function in the context of procedural operations in NARS.

Importantly though, the study by Hayes and others is only one among several hundreds of well-conducted studies investigating different aspects of RFT. Many of these studies could be used for creating research programs in the intersection of experimental psychology and NARS.

RFT was developed in the late 1980's with clinical applications in mind [1]. There are many studies conducted on how the processes described in this text are related to human suffering. RFT assumes that the SELF-concept is at the root of these clinical processes. More specifically, a person's conceptual SELF-image that is generated over time is tied into an endless amount of relational networks. People put themselves in comparison and opposition to others, and functions of the SELF are transformed accordingly. Given the centrality of the SELF in NARS, and how it is related to other concepts over time, there seems to be a natural parallel between NARS and clinical applications grounded in RFT. A lot of future research opportunities seem to be available in this domain.

9 Summary

In summary, NARS seems promising for studying AARR in machines. Some relations are built-in into NARS, and others can be learned from experience. Mutually entailed responses seem to work well, while there is work left to make NARS reliably produce combinatorially entailed relational responses. Simple stimulus functions have been explored and we have showed above how these can be transformed through various relational networks. Future challenges include demonstrating relational responding as part of procedural operations in NARS.

Acknowledgements The author would like to thank the NARS team for the very helpful discussions. A special acknowledgement to Tony Lofthouse who has taken a lot of his time to discuss NARS and RFT.

References

1. Hayes, S.C., Barnes-Holmes, D., Roche, B.: Relational frame theory : a post-Skinnerian account of human language and cognition. Kluwer Academic/Plenum Publishers, New York (2001)
2. Hayes, S.C., Devany, J.M., Kohlenberg, B.S., Brownstein, A.J., Shelby, J.: Stimulus equivalence and the symbolic control of behavior. *Revista Mexicana de Analisis de la Conducta* **13**(3), 361–374 (1987)
3. Johansson, R.: Arbitrarily applicable relational responding. In: *International Conference on Artificial General Intelligence*. Springer (2019)
4. Sidman, M.: Equivalence relations and behavior: An introductory tutorial. *The Analysis of verbal behavior* **25**(1), 5–17 (2009)
5. Steele, D., Hayes, S.C.: Stimulus equivalence and arbitrarily applicable relational responding. *Journal of the Experimental Analysis of Behavior* **56**(3), 519–555 (1991)
6. Wang, P., Hammer, P.: Perception from an agi perspective. In: *International Conference on Artificial General Intelligence*. pp. 259–269. Springer (2018)
7. Wang, P., Li, X., Hammer, P.: Self in nars, an agi system. *Frontiers in Robotics and AI* **5**, 20 (2018)