An improved upper bound for the number of intermediate inputs (unpublished draft)

Alexander Nittka

July 8, 2008

1 Introduction

When allowing an observation o to be partial in the sense that additional (unrecorded) inputs may have been received during the time of observation, we showed that reasoning about the agent is possible by adding intermediate inputs to the observation, instantiate them by new variables yielding an observation o' in the traditional sense. The rational explanation of o' then provides a basis for reasoning about $\mathcal{A}$. The question then is how many intermediate inputs to assume at which positions. We showed that an additional intermediate input may cause the weakest acceptable core belief to become yet weaker. Consequently, we want to assume enough intermediate inputs in order to get the overall weakest acceptable core in order to draw safe conclusions as to which revision inputs are rejected by the agent.

Here we are interested in the case where the positions of the intermediate inputs are known. That is, we know when in the observation additional inputs were received but not how many. In [1] we provided first results on an upper bound. This paper improves these results. In particular it reduces the bound given by Proposition 3.26 (upperboundii). That proposition entails that the number of intermediate inputs that need to be assumed at a certain position in the observation is bounded by the number of all recorded revision inputs following that position. We will show that the number of additional inputs at one position need not exceed the number of recorded inputs following that position *but before the next block of intermediate inputs*.

The number of intermediate inputs is thus always less than the number of recorded inputs. As a rule of thumb we could say that one intermediate input per recorded one suffices. Using Proposition 3.26 this number could be much greater. Consider one hundred positions for intermediate inputs near the beginning of an observation followed by one million recorded inputs. The original result would call for about a hundred million intermediate inputs. The improved version reduces that number to about one million.

As a consequence, the result presented here also generalises Proposition 3.23 (oneiienough) in [1] which states that being allowed an arbitrary number of intermediate inputs at any position in the observation, a single input between any two recorded inputs suffices. There is only one recorded input between any two consecutive positions for intermediate inputs.

2 Basic Idea

The actual revision history of the agent whose core belief is ▲ looks like this:

$$(\alpha_1, \dots, \alpha_k, \varphi_{01}, \dots, \varphi_{0n_0}, \psi_{11}, \dots, \psi_{1m_1}, \varphi_{11}, \dots, \varphi_{1n_1}, \psi_{21}, \dots, \psi_{2m_2}, \dots, \varphi_{l1}, \dots, \varphi_{ln_l})$$

The observation starts at the point where φ_{01} was received and contains the recorded inputs $\varphi_{01}, \dots, \varphi_{0n_0}, \varphi_{11}, \dots, \varphi_{1n_1}, \dots, \varphi_{l1}, \dots, \varphi_{ln_l}$ but does not contain the inputs ψ_{ij}. These have to be dealt with using intermediate inputs. $\alpha_1, \dots, \alpha_k$ are the formulae received by $\mathcal{A}$ before the observation started.

The idea is to replace each block of unrecorded inputs $\psi_{i1}, \dots, \psi_{im_i}$ by (at most) n_i formulae such that the beliefs and non-beliefs for all recorded revision inputs in the observation stay the same. Note that n_i is the number of recorded revision inputs following that block of intermediate inputs before the next block of intermediate inputs starts. As we assume to know the positions of the intermediate inputs, we know all n_j but not the m_j.

Each block $\psi_{i1}, \dots, \psi_{im_i}$ will be replaced by a logical chain $\lambda_{i1}, \dots, \lambda_{in_i}$,[1] the idea being that each of the recorded inputs φ_{ij} immediately following that block will select one element of the chain such that the exactly the same beliefs are generated *from that chain alone.* That is, revision inputs recorded (or not recorded) earlier are made redundant. We will replace each block independently of the others. $\mathcal{A}$'s revision history and the modified one, which yields the same beliefs for all recorded revision inputs, look as follows.

$$(\alpha_1, \dots, \alpha_k, \varphi_{01}, \dots, \varphi_{0n_0}, \psi_{11}, \dots, \psi_{1m_1}, \varphi_{11}, \dots, \varphi_{1n_1}, \psi_{21}, \dots, \psi_{2m_2}, \dots, \varphi_{l1}, \dots, \varphi_{ln_l})$$
$$(\alpha_1, \dots, \alpha_k, \varphi_{01}, \dots, \varphi_{0n_0}, \lambda_{11}, \dots, \lambda_{1n_1}, \varphi_{11}, \dots, \varphi_{1n_1}, \lambda_{21}, \dots, \lambda_{2n_2}, \dots, \varphi_{l1}, \dots, \varphi_{ln_l})$$

Two things will need to be shown: (i) such logical chains do exist and (ii) replacing a block by a logical chain will have no negative effect for beliefs of later blocks. Note that when calculating the beliefs using f, not the original formulae of the revision history are used, but the formulae from the logical chains.

[1] Note that m_i could be much greater than n_i and that the length of this logical chain corresponds to the claim that one intermediate input per recorded one before a new block of intermediate inputs.

The structure of the λ_{ij} will be given in Section 4 along with the proofs of the existence of the logical chains and that the beliefs are indeed the same. But what do we win by this transformation? We have an observation of $\mathcal{A}$ which is based on its revision history. We know the positions of the intermediate inputs, i.e. we know all $n_1, \ldots, n_l$, but we do not know the number of intermediate inputs, i.e. we do not know any m_j. The transformation now tells us that there are intermediate inputs (the λ_{ij} of which we know number and positions!) that yield exactly the same beliefs for each recorded input. Applying Proposition 3.6 (nxworksgen) in [1] tells us that we can replace these unknown intermediate inputs by new propositional variables and be guaranteed that an explanation $[\rho_R, \blacktriangle']$ is found. Propositions 3.12 (implynxcoregen) and 3.13 (weakestnxcoregen) in [1] further yield that $\mathcal{A}$'s core belief must entail $\blacktriangle'$. We can also use the explanation to do hypothetical reasoning about $\mathcal{A}$'s beliefs. However, note that for reasoning about future beliefs additional intermediate inputs must be added to the last block of intermediate inputs. Otherwise there would be a violation of the assumption that there is an intermediate input for each recorded one following that block.

3 Auxiliary Results

The following is a generalisation of Proposition 2.15 (entailprefix) in [1].

Proposition 1. *Either* $f(\sigma \cdot \rho_1 \cdot \rho_2 \cdot \alpha) \vdash \neg f(\rho_1 \cdot \alpha)$ *or* $f(\sigma \cdot \rho_1 \cdot \rho_2 \cdot \alpha) \vdash f(\rho_1 \cdot \alpha)$

Proof. $f(\sigma \cdot \rho_1 \cdot \rho_2 \cdot \alpha) \equiv f(\sigma \cdot f(\rho_1 \cdot \rho_2 \cdot \alpha))$ (Proposition 2.3 (fseqtopair) in [1]). By definition of f this formula entails $f(\rho_1 \cdot \rho_2 \cdot \alpha)$ and by Proposition 2.15 (entailprefix) in [1] this formula in turn entails either $\neg f(\rho_1 \cdot \alpha)$ or $f(\rho_1 \cdot \alpha)$. $\square$

The following proposition will help us to show that the encodings we use for the intermediate inputs do not have any negative effects.

Proposition 2. *Let* $\sigma = (\beta_1, \ldots, \beta_m)$ *be some sequence of formulae and* $\blacktriangle$ *a formula. Let* ${}_i\sigma_j$ *denote* $(\beta_i, \ldots, \beta_j)$ *for all* $i \leq j$ *and let* σ_j *denote* $(\beta_1, \ldots, \beta_j)$*. Then for all* $i \leq j \leq k \leq m$*:* $f(\sigma_k \cdot \blacktriangle) \vdash f({}_i\sigma_j \cdot \blacktriangle) \rightarrow f(\sigma_j \cdot \blacktriangle)$ *or* $f(\sigma_k \cdot \blacktriangle)$ *is inconsistent with* $f({}_i\sigma_j \cdot \blacktriangle) \rightarrow f(\sigma_j \cdot \blacktriangle)$*.*

Proof. Follows almost directly from Proposition 2.15 (entailprefix) in [1] and Proposition 1. Let $\sigma \cdot \rho_1 \cdot \rho_2 = \sigma_k$, $\rho_1 = {}_i\sigma_j$ and $\sigma \cdot \rho_1 = \sigma_j$. Then we immediately get $f(\sigma_k \cdot \blacktriangle) \vdash f({}_i\sigma_j \cdot \blacktriangle)$ or $f(\sigma_k \cdot \blacktriangle) \vdash \neg f({}_i\sigma_j \cdot \blacktriangle)$.

Analogously Propositions 2.15 (entailprefix) in [1] yields $f(\sigma_k \cdot \blacktriangle) \vdash f(\sigma_j \cdot \blacktriangle)$ or $f(\sigma_k \cdot \blacktriangle) \vdash \neg f(\sigma_j \cdot \blacktriangle)$. And for any combination $f(\sigma_k \cdot \blacktriangle)$ either entails the implication or is inconsistent with it. $\square$

4 Main Results

Let us start by describing the principal structure of a block of replaced intermediate inputs. Focussing on the part that interests us — a single block of intermediate inputs and the recorded revision inputs immediately following — the true and the modified revision history of $\mathcal{A}$ are: $(\ldots, \psi_{i1}, \ldots, \psi_{im_i}, \varphi_{i1}, \ldots, \varphi_{in_i}, \ldots)$ and $(\ldots, \lambda_{i1}, \ldots, \lambda_{in_i}, \varphi_{i1}, \ldots, \varphi_{in_i}, \ldots)$. Let σ denote $(\ldots, \psi_{i1}, \ldots, \psi_{im_i})$. Let ι denote $(\varphi_{i1}, \ldots, \varphi_{in_1})$ and ι_j the prefix of that sequence with length j, i.e. $\iota_j = (\varphi_{i1}, \ldots, \varphi_{ij})$. Further, let $S = \{1, \ldots, n_i\}$. ι is the sequence of recorded revision inputs for which we want to generate the same beliefs, σ is the sequence of all inputs received up to that point. For readability we will omit the first of the double indices and write λ_j instead of λ_{ij}, φ_n instead of φ_{in_i} etc. Thus i will be used as a regular variable from now on.

Each λ_j will have the following form: $\lambda_j = \bigwedge\limits_{i \in S_j} (f(\iota_i \cdot \blacktriangle) \rightarrow f(\sigma \cdot \iota_i \cdot \blacktriangle))$ for a suitable $S_j \subseteq S$. We demand that $S_n = S$, that $S_k \subseteq S_j$ for all $k \leq j$ and that $i \in S_{j-1}$ in case $f(\iota_i \cdot \blacktriangle)$ happens to be inconsistent with λ_j. We will later show that these demands can be met and first try to give an intuition of what the λ_j mean.

Enforcing the subset relation among the S_j amounts to $(\lambda_1, \ldots, \lambda_n)$ being a logical chain. The beliefs after receiving the revision input φ_i are $f(\sigma \cdot (\varphi_1, \ldots, \varphi_i, \blacktriangle)) = f(\sigma \cdot \iota_i \cdot \blacktriangle) = f(\sigma \cdot f(\iota_i \cdot \blacktriangle))$. The suffix of σ is $(\psi_{i1}, \ldots, \psi_{im_i})$. For the modified revision history it will be analogous for some sequence σ'. Note that the suffix of σ' will be the logical chain $(\lambda_1, \ldots, \lambda_n)$. Before going through the logical chain the formula $f(\iota_i \cdot \blacktriangle)$ will have been collected. Then exactly one λ_j is selected — $f(\iota_i \cdot \blacktriangle)$ will be inconsistent with the λ_k, $k > j$ and those for $k < j$ can be ignored as they are already entailed. Due to our demands $i \in S_j$ and thus $\lambda_j \vdash f(\iota_i \cdot \blacktriangle) \rightarrow f(\sigma \cdot \iota_i \cdot \blacktriangle)$. Now, $f(\iota_i \cdot \blacktriangle) \wedge (f(\iota_i \cdot \blacktriangle) \rightarrow f(\sigma \cdot \iota_i \cdot \blacktriangle)) \vdash f(\sigma \cdot \iota_i \cdot \blacktriangle)$. That is, after processing the logical chain, the formula collected so far already entails the beliefs we are after.

We still need to show that (i) after processing the logical chain, *nothing but* $f(\sigma \cdot \iota_i \cdot \blacktriangle)$ is entailed, i.e. $f((\lambda_1, \ldots, \lambda_n) \cdot \iota_i \cdot \blacktriangle) \equiv f(\sigma \cdot \iota_i \cdot \blacktriangle)$, that (ii) going through the rest of σ' does not add any further beliefs and (iii) that a logical chain satisfying the demand mentioned above does indeed exist.

4.1 $f((\lambda_1, \ldots, \lambda_n) \cdot \iota_i \cdot \blacktriangle) \equiv f(\sigma \cdot \iota_i \cdot \blacktriangle)$

As mentioned before, exactly one λ_j is selected from $(\lambda_1, \ldots, \lambda_n)$. The others are either inconsistent with $f(\iota_i \cdot \blacktriangle)$ or already entailed by λ_j. Consequently $f((\lambda_1, \ldots, \lambda_n) \cdot \iota_i \cdot \blacktriangle) \equiv \lambda_j \wedge f(\iota_i \cdot \blacktriangle)$ and we need to show $\lambda_j \wedge f(\iota_i \cdot \blacktriangle) \equiv f(\sigma \cdot \iota_i \cdot \blacktriangle)$. We already showed that

$\lambda_j \vdash f(\iota_i \cdot \blacktriangle) \to f(\sigma \cdot \iota_i \cdot \blacktriangle)$ and thus $\lambda_j \wedge f(\iota_i \cdot \blacktriangle) \vdash f(\sigma \cdot \iota_i \cdot \blacktriangle)$. We still need to show $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \lambda_j \wedge f(\iota_i \cdot \blacktriangle)$. We will do so by showing that nothing but $f(\sigma \cdot \iota_i \cdot \blacktriangle)$ can be inferred from $\lambda_j \wedge f(\iota_i \cdot \blacktriangle)$.

By the structure of λ_j, we know $\lambda_j \wedge f(\iota_i \cdot \blacktriangle) \equiv f(\sigma \cdot \iota_i \cdot \blacktriangle) \wedge \bigwedge_{k \in S_j \setminus \{i\}} (f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle))$
For $k < i$ we immediately have that $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle)$ (Propostion 2; $\sigma_k = \sigma \cdot \iota_i$, $\sigma_j = \sigma \cdot \iota_k$, ${}_i\sigma_j = \iota_k$; it cannot be inconsistent with the implication as $\lambda_j \wedge f(\iota_i \cdot \blacktriangle)$ is consistent). So let $k > i$.

- $f(\iota_i \cdot \blacktriangle)$ is inconsistent with $f(\iota_k \cdot \blacktriangle)$ implying $f(\iota_i \cdot \blacktriangle) \vdash f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle))$.
- $f(\iota_i \cdot \blacktriangle)$ is consistent with $f(\iota_k \cdot \blacktriangle)$. By Proposition 2.15 (entailprefix) in [1] $f(\iota_k \cdot \blacktriangle) \vdash f(\iota_i \cdot \blacktriangle)$ and they select the same elements from ι_i.
  - $f(\sigma \cdot \iota_i \cdot \blacktriangle)$ and $f(\sigma \cdot \iota_k \cdot \blacktriangle)$ select exactly the same elements from σ. Let χ denote the conjunction of those elements selected from σ, so $f(\sigma \cdot \iota_i \cdot \blacktriangle) \equiv \chi \wedge f(\iota_i \cdot \blacktriangle)$ and $f(\sigma \cdot \iota_k \cdot \blacktriangle) \equiv \chi \wedge f(\iota_k \cdot \blacktriangle)$.
    Consequently $f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle)) \equiv f(\iota_k \cdot \blacktriangle) \to \chi \wedge f(\iota_k \cdot \blacktriangle)$. This is equivalent to $(f(\iota_k \cdot \blacktriangle) \to \chi) \wedge (f(\iota_k \cdot \blacktriangle) \to f(\iota_k \cdot \blacktriangle))$.
    Now, $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \chi$ and $f(\iota_k \cdot \blacktriangle) \to f(\iota_k \cdot \blacktriangle)$ is a tautology and thus we have $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle)$.
  - $f(\sigma \cdot \iota_i \cdot \blacktriangle)$ and $f(\sigma \cdot \iota_k \cdot \blacktriangle)$ select different elements from σ. This means there is at least one formula α in σ such that $[f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \alpha$ and $f(\sigma \cdot \iota_k \cdot \blacktriangle) \vdash \neg\alpha]$ or $[f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \neg\alpha$ and $f(\sigma \cdot \iota_k \cdot \blacktriangle) \vdash \alpha]$.
    $\lambda_j \vdash f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle)$ and thus $\lambda_j \wedge f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \alpha \wedge (f(\iota_k \cdot \blacktriangle) \to \neg\alpha)$ and hence $\lambda_j \wedge f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \neg f(\iota_k \cdot \blacktriangle)$. (Analogous for $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \neg\alpha$.) If we can show that $f(\sigma \cdot \iota_i \cdot \blacktriangle)$ alone already entails $\neg f(\iota_k \cdot \blacktriangle)$ we immediately have $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle)$.
    So let $\sigma = \sigma'' \cdot \alpha \cdot \sigma'$ where α is the first element in σ for which $f(\sigma \cdot \iota_i \cdot \blacktriangle)$ and $f(\sigma \cdot \iota_k \cdot \blacktriangle)$ differ, i.e. both select the same elements from σ' (and as shown above also the same elements from ι_i). $f(\alpha \cdot \sigma' \cdot \iota_i \cdot \blacktriangle) \equiv f(\alpha \cdot f(\sigma' \cdot \iota_i \cdot \blacktriangle))$ and $f(\alpha \cdot \sigma' \cdot \iota_k \cdot \blacktriangle) \equiv f(\alpha \cdot f(\sigma' \cdot \iota_k \cdot \blacktriangle))$.
    $f(\alpha \cdot f(\sigma' \cdot \iota_i \cdot \blacktriangle))$ accepts α and $f(\alpha \cdot f(\sigma' \cdot \iota_k \cdot \blacktriangle))$ rejects α. (Assume that $f(\alpha \cdot f(\sigma' \cdot \iota_i \cdot \blacktriangle))$ rejects α. Then $f(\sigma' \cdot \iota_i \cdot \blacktriangle) \vdash \neg\alpha$, but $f(\iota_k \cdot \blacktriangle) \vdash f(\iota_i \cdot \blacktriangle)$ and they select the same elements from σ' so $f(\sigma' \cdot \iota_k \cdot \blacktriangle) \vdash \neg\alpha$ — contradiction as they are supposed to behave differently for α.)
    Now let us look more closely at the structure of $f(\sigma' \cdot \iota_i \cdot \blacktriangle)$ and $f(\sigma' \cdot \iota_k \cdot \blacktriangle)$. Let χ_1 denote the conjunction of elements selected from σ', let χ_2 denote the conjunction of elements selected from $\iota_i = (\varphi_1, \ldots, \varphi_i)$ and let β denote the elements

selected from ${}_{(i+1)}\iota_k = (\varphi_{i+1}, \ldots, \varphi_k)$. That is $f(\sigma' \cdot \iota_i \cdot \blacktriangle) \equiv \blacktriangle \wedge \chi_1 \wedge \chi_2$ and $f(\sigma' \cdot \iota_k \cdot \blacktriangle) \equiv \blacktriangle \wedge \chi_1 \wedge \chi_2 \wedge \beta$.

$f(\sigma' \cdot \iota_k \cdot \blacktriangle)$ rejects α. Hence $\blacktriangle \wedge \chi_1 \wedge \chi_2 \wedge \beta \vdash \neg\alpha$ implying $\blacktriangle \wedge \chi_1 \wedge \chi_2 \wedge \alpha \vdash \neg\beta$. This yields $\blacktriangle \wedge \chi_1 \wedge \chi_2 \wedge \alpha \vdash \neg\blacktriangle \vee \neg\chi_2 \vee \neg\beta$ which translates directly into $f(\alpha \cdot f(\sigma' \cdot \iota_i \cdot \blacktriangle)) \vdash \neg f(\iota_k \cdot \blacktriangle)$. Thus we have $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \neg f(\iota_k \cdot \blacktriangle)$ which we wanted to show.

We have thus shown $f(\sigma \cdot \iota_i \cdot \blacktriangle) \vdash \bigwedge_{k \in S_j \setminus \{i\}} (f(\iota_k \cdot \blacktriangle) \to f(\sigma \cdot \iota_k \cdot \blacktriangle))$. We also know that $(f(\iota_i \cdot \blacktriangle) \to f(\sigma \cdot \iota_i \cdot \blacktriangle)) \wedge f(\iota_i \cdot \blacktriangle) \equiv f(\sigma \cdot \iota_i \cdot \blacktriangle)$ and hence $\lambda_j \wedge f(\iota_i \cdot \blacktriangle) \equiv f(\sigma \cdot \iota_i \cdot \blacktriangle)$.

4.2 No new beliefs when processing the rest of σ'

Let us sum up what we know so far. We have an agent's revision history. We have replaced *each* block of intermediate inputs by a logical chain of the form described before. All other elements of the revision history remain unchanged. That is, beliefs before any block of intermediate inputs remain unchanged — they are the same for the original and the modified revision history.

$$
\begin{array}{lllll}
\sigma^i & \iota^i & \sigma^j & & \iota^j \\
(\ldots, \psi_{i1}, \ldots, \psi_{im_i}, & \varphi_{i1}, \ldots, \varphi_{in_i}, & \psi_{(i+1)1}, \ldots, \varphi_{(j-1)n_{j-1}}, & \psi_{j1}, \ldots, \psi_{jm_j}, & \varphi_{j1}, \ldots, \varphi_{jn_j}) \\
(\ldots, \lambda_{i1}, \ldots, \lambda_{in_i}, & \varphi_{i1}, \ldots, \varphi_{in_i}, & \lambda_{(i+1)1}, \ldots, \varphi_{(j-1)n_{j-1}}, & \lambda_{j1}, \ldots, \lambda_{jn_j}, & \varphi_{j1}, \ldots, \varphi_{jn_j})
\end{array}
$$

For recorded revision inputs after some block of intermediate inputs the previous section told us the following. When calculating the agent's beliefs after having received a recorded revision input φ_{jk}, we go through the immediately preceding block of replaced intermediate inputs $\lambda_{j1}, \ldots, \lambda_{jn_j}$. At that point we already have exactly the same beliefs as the agent had for the original revision history, i.e. $f\left(\sigma^i \cdot \iota^i \cdot \sigma^j \cdot \iota^j_k \cdot \blacktriangle\right)$.

Now we need to show that no further beliefs are introduced by processing the rest of the modified sequence. All elements that remained unchanged are either already entailed or inconsistent with the beliefs collected so far. So we only have to make sure that the elements λ of all the other logical chains before the one that has been processed already don't do any harm. So let us consider an arbitrary λ_{il}, $i < j$ and $1 \leq l \leq n_i$.

By construction $\lambda_{il} \equiv \bigwedge_{s \in S} (f(\iota^i_s \cdot \blacktriangle) \to f(\sigma^i \cdot \iota^i_s \cdot \blacktriangle))$ for a suitable $S \subseteq \{1, \ldots, n_i\}$. If we can show that $f\left(\sigma^i \cdot \iota^i \cdot \sigma^j \cdot \iota^j_k \cdot \blacktriangle\right)$ entails all or is inconsistent with any of these implications, we know that λ_{il} introduces no further belief as it is already entailed or is inconsistent with what has been collected so far. Note that $\iota^i = \iota^i_s \cdot {}_{s+1}\iota^i_{n_i}$ (for $1 \leq s < n_i$) and $\iota^i = \iota^i_s$ (for $s = n_i$). We will consider only the first case; the second one is analgous.

It holds that $f\left(\sigma^i \cdot \iota^i \cdot \sigma^j \cdot \iota^j_k \cdot \blacktriangle\right) = f\left(\sigma^i \cdot \iota^i_s \cdot {}_{s+1}\iota^i_{n_i} \cdot \sigma^j \cdot \iota^j_k \cdot \blacktriangle\right)$ entails $f\left(\iota^i_s \cdot \blacktriangle\right)$ or is inconsistent with it. This is due to Proposition 1 using the following substitutions: $\sigma = \sigma^i$, $\rho_1 = \iota^i_s$, $\rho_2 = {}_{s+1}\iota^i_{n_i} \cdot \sigma^j \cdot \iota^j_k$, $\alpha = \blacktriangle$. Also $f\left(\sigma^i \cdot \iota^i_s \cdot {}_{s+1}\iota^i_{n_i} \cdot \sigma^j \cdot \iota^j_k \cdot \blacktriangle\right)$ entails $f\left(\sigma^i \cdot \iota^i_s \cdot \blacktriangle\right)$ or is inconsistent with it. This is due to Proposition 2.15 (entailprefix) in [1] using the following substitutions: $\sigma = \sigma^i \cdot \iota^i_s$, $\rho = {}_{s+1}\iota^i_{n_i} \cdot \sigma^j \cdot \iota^j_k$, $\alpha = \blacktriangle$. Consequently, for any combination of these cases $f\left(\sigma^i \cdot \iota^i \cdot \sigma^j \cdot \iota^j_k \cdot \blacktriangle\right)$ entails the implication $f\left(\iota^i_s \cdot \blacktriangle\right) \rightarrow f\left(\sigma^i \cdot \iota^i_s \cdot \blacktriangle\right)$ or is inconsistent with it. Hence λ_{il} is either entailed or inconsistent with what has been collected so far. In both cases λ does not modify the beliefs.

4.3 The existence of the logical chains

The final step is the proof of the existence of logical chains with the properties we assumed. Recall the structure of the original revision history and the modified one, focussing a single block of intermediate inputs to be replaced: $(\ldots, \psi_1, \ldots, \psi_m, \varphi_1, \ldots, \varphi_n, \ldots)$ and $(\ldots, \lambda_1, \ldots, \lambda_n, \varphi_1, \ldots, \varphi_n, \ldots)$. The agent's core belief $\blacktriangle$ is known. Let σ denote $(\ldots, \psi_1, \ldots, \psi_m)$ and let ι denote $(\varphi_1, \ldots, \varphi_n)$. Our goal is that $f(\sigma \cdot \iota_j \cdot \blacktriangle)$ is believed after φ_j has been received and some logical chain has been processed. Now the rational explanation is just the tool for constructing a logical chain. We simply translate our requirement into the observation $o = \langle(\varphi_1, f(\sigma \cdot \iota_1 \cdot \blacktriangle), \emptyset), \ldots, (\varphi_n, f(\sigma \cdot \iota_n \cdot \blacktriangle), \emptyset)\rangle$. Recall that the conditional beliefs corresponding to this observation are $f((\varphi_1, \ldots, \varphi_i) \cdot \blacktriangle) \Rightarrow f(\sigma \cdot \iota_i \cdot \blacktriangle)$, the antecedent being nothing but $f(\iota_i \cdot \blacktriangle)$.

Now we know that this observation has an explanation: $[\sigma, \blacktriangle]$. Hence $\blacktriangle$ is o-acceptable and the rational prefix construction using $\blacktriangle$ as the initial core belief must find a suitable logical chain (Propositions 2.49 and 2.50 (ratprefok and ratprefnotok) in [1]). The rational prefix construction uses the material counterparts $f(\iota_i \cdot \blacktriangle) \rightarrow f(\sigma \cdot \iota_i \cdot \blacktriangle)$ of the conditionals. Note that the resulting implications have exactly the form we proposed for constructing the λ_j. As there are no non-beliefs in the observation each element in the logical chain constructed will be a conjunction of such implications. How many elements will the logical chain have? In every iteration not all of the remaining conditionals can be p-exceptional (this would imply that $\blacktriangle$ is not o-acceptable which it is), at least one conditional is eliminated in each step. There are n conditionals which yields at most $n+1$ sets. However the last one will be empty, the corresponding formula being a tautology. Hence there are at most n non-trivial elements in the logical chain.[2] And they have exactly the form proposed for the λ_j.

[2] There could be fewer than n elements in which case we can add tautologies without negative impact.

5 Conclusion

We have thus shown that (in principal) we can use the rational closure to replace each block of intermediate inputs by a block of the proposed length. For this the entire revision history would have to be known. Of course we know neither the core belief nor the revision history — otherwise we would not need complicated tools for reasoning about the agent's beliefs, we could simply calculate them. This paper merely proves the *existence* of such logical chains. This yields an improved bound on the number of intermediate inputs that need to be assumed.

References

[1] Alexander Nittka. *A Method for Reasoning About other Agents' Beliefs from Observations*. PhD thesis, Leipzig University, 2008.