

Verification of Discrete Upper Confidence Bound algorithm in Isabelle/HOL

Arjan Faber

February 6, 2026

Abstract

This project formally verifies the Upper Confidence Bound (UCB) algorithm in Isabelle/Higher-order Logic (HOL), focusing on its probabilistic guarantees and regret bounds. The work extends Isabelle/HOLs probabilistic framework and explores verification of discrete-time bandit models following [1]. This research advances the formal verification of probabilistic algorithms in reinforcement learning.

theory *MSc-Project-Discrete-Prop15-1*

imports

HOL-Probability.Probability

begin

locale *bandit-problem* =

fixes $A :: 'a \text{ set}$

and $\mu :: 'a \Rightarrow \text{real}$

and $a\text{-star} :: 'a$

assumes *finite-arms*: $\text{finite } A$

and *a-star-in-A*: $a\text{-star} \in A$

and *optimal-arm*: $\forall a \in A. \mu a\text{-star} \geq \mu a$

begin

definition $\Delta :: 'a \Rightarrow \text{real}$ **where**

$\Delta a = \mu a\text{-star} - \mu a$

end

locale *bandit-policy* = *bandit-problem* + *prob-space* +

fixes $\Omega :: 'b \text{ set}$

and $\mathcal{F} :: 'b \text{ set set}$

and $\pi :: \text{nat} \Rightarrow 'b \Rightarrow 'a$

and $N\text{-}n :: \text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{nat}$

assumes *measurable-policy*: $\forall t. \pi t \in \text{measurable } M \text{ (count-space } A)$

and *N-n-def*: $\forall n a \omega. N\text{-}n n a \omega = \text{card } \{t \in \{0..<n\}. \pi (t+1) \omega = a\}$

and *count-assm-pointwise*: $\forall n \omega. (\sum a \in A. \text{real } (N\text{-}n n a \omega)) = \text{real } n$

begin

definition $R\text{-}n :: \text{nat} \Rightarrow 'b \Rightarrow \text{real}$ **where**

$$R\text{-}n\ n\ \omega = n * \mu\ a\text{-}star - (\sum a \in A. \mu\ a * \text{real}\ (N\text{-}n\ n\ a\ \omega))$$

lemma *regret-decomposition-pointwise*:

fixes $n :: \text{nat}$ **and** $\omega :: 'b$

assumes $n\text{-count-asm-pointwise}$: $(\sum a \in A. \text{real}\ (N\text{-}n\ n\ a\ \omega)) = \text{real}\ n$

shows $R\text{-}n\ n\ \omega = (\sum a \in A. \Delta\ a * \text{real}\ (N\text{-}n\ n\ a\ \omega))$

proof –

have $sum\text{-}Nn$: $(\sum a \in A. \text{real}\ (N\text{-}n\ n\ a\ \omega)) = \text{real}\ n$

using $n\text{-count-asm-pointwise}$ **by** $simp$

have $sum\text{-}const$: $(\sum a \in A. \mu\ a\text{-}star * \text{real}\ (N\text{-}n\ n\ a\ \omega)) = \mu\ a\text{-}star * (\sum a \in A. \text{real}\ (N\text{-}n\ n\ a\ \omega))$

by ($simp\ add$: $Cartesian\text{-}Space.vector\text{-}space\text{-}over\text{-}itself.scale\text{-}sum\text{-}right$)

have $eq1$: $R\text{-}n\ n\ \omega = \text{real}\ n * \mu\ a\text{-}star - (\sum a \in A. \mu\ a * \text{real}\ (N\text{-}n\ n\ a\ \omega))$

by ($simp\ add$: $R\text{-}n\text{-}def$)

also have $eq2$: $\dots = (\sum a \in A. \mu\ a\text{-}star * \text{real}\ (N\text{-}n\ n\ a\ \omega)) - (\sum a \in A. \mu\ a * \text{real}\ (N\text{-}n\ n\ a\ \omega))$

using $sum\text{-}const\ sum\text{-}Nn$

by ($subst\ sum\text{-}Nn[symmetric]$, $subst\ sum\text{-}const$) $simp$

also have $eq3$: $\dots = (\sum a \in A. (\mu\ a\text{-}star * \text{real}\ (N\text{-}n\ n\ a\ \omega) - \mu\ a * \text{real}\ (N\text{-}n\ n\ a\ \omega)))$

by ($rule\ sum\text{-}subtractf[symmetric]$)

also have $eq4$: $\dots = (\sum a \in A. (\mu\ a\text{-}star - \mu\ a) * \text{real}\ (N\text{-}n\ n\ a\ \omega))$

by ($simp\ add$: $algebra\text{-}simps$)

also have $\dots = (\sum a \in A. \Delta\ a * \text{real}\ (N\text{-}n\ n\ a\ \omega))$

by ($simp\ add$: $\Delta\text{-}def$)

finally show $?thesis$.

qed

lemma *integrable-const-fun*:

assumes $finite\text{-}measure\ M$

shows $integrable\ M\ (\lambda x. c)$

using $assms$ **by** ($simp\ add$: $Bochner\text{-}Integration.finite\text{-}measure.integrable\text{-}const$)

lemma *expected-regret*:

assumes $finite\ A$

and $\forall a \in A. integrable\ M\ (\lambda \omega. \text{real}\ (N\text{-}n\ n\ a\ \omega))$

shows $expectation\ (\lambda \omega. R\text{-}n\ n\ \omega) = (\sum a \in A. \Delta\ a * expectation\ (\lambda \omega. \text{real}\ (N\text{-}n\ n\ a\ \omega)))$

proof –

have *integrable-sum*: *integrable* M ($\lambda\omega. \sum a \in A. \mu a * \text{real } (N\text{-}n \ n \ a \ \omega)$)

proof –

have $\forall a \in A. \text{integrable } M$ ($\lambda\omega. \mu a * \text{real } (N\text{-}n \ n \ a \ \omega)$)

using *assms integrable-mult-right* **by** *blast*

then show *?thesis*

by (*simp add: integrable-sum*)

qed

have *pointwise*: $\forall\omega. R\text{-}n \ n \ \omega = (\sum a \in A. \Delta a * \text{real } (N\text{-}n \ n \ a \ \omega))$

using *regret-decomposition-pointwise count-assm-pointwise* **by** *blast*

hence *rewrite*: *expectation* ($\lambda\omega. R\text{-}n \ n \ \omega$) = *expectation* ($\lambda\omega. \sum a \in A. \Delta a * \text{real } (N\text{-}n \ n \ a \ \omega)$)

by *simp*

also have *expectation* ($\lambda\omega. R\text{-}n \ n \ \omega$) = ($\sum a \in A. \Delta a * \text{expectation } (\lambda\omega. \text{real } (N\text{-}n \ n \ a \ \omega))$)

proof –

have $\forall a \in A. \text{integrable } M$ ($\lambda\omega. \Delta a * \text{real } (N\text{-}n \ n \ a \ \omega)$)

using *assms integrable-mult-right* **by** *blast*

hence *integrable* M ($\lambda\omega. \sum a \in A. \Delta a * \text{real } (N\text{-}n \ n \ a \ \omega)$)

using *integrable-sum* **by** *simp*

have *res-1*: *expectation* ($\lambda\omega. \sum a \in A. \Delta a * \text{real } (N\text{-}n \ n \ a \ \omega)$) = *integral^L* M ($\lambda\omega. \sum a \in A. \Delta a * \text{real } (N\text{-}n \ n \ a \ \omega)$)

by *simp*

also have *res-2*: $\dots = (\sum a \in A. \text{integral^L} \ M$ ($\lambda\omega. \Delta a * \text{real } (N\text{-}n \ n \ a \ \omega)$))

using *<finite A> assms*

by (*simp add: integral-sum integrable-sum*)

have *res-3*: *expectation* ($\lambda\omega. R\text{-}n \ n \ \omega$) = ($\sum a \in A. \Delta a * \text{expectation } (\lambda\omega. \text{real } (N\text{-}n \ n \ a \ \omega))$)

using *assms rewrite*

by *simp*

then show *?thesis* **using** *assms res-3* **by** *auto*

qed

then show *?thesis* .

qed

end

```

end
theory Discrete-UCB-Step1
  imports MSc-Project-Discrete-Prop15-1

begin

locale bandit-policy = bandit-problem + prob-space +
  fixes  $\Omega$  :: 'b set
  and  $\mathcal{F}$  :: 'b set set
  and  $\omega$  :: 'b
  and  $\pi$  :: nat  $\Rightarrow$  'b  $\Rightarrow$  'a
  and  $N$ -n :: nat  $\Rightarrow$  'a  $\Rightarrow$  'b  $\Rightarrow$  nat
  assumes measurable-policy:  $\forall t. \pi t \in \text{measurable } M \text{ (count-space } A)$ 
  and  $N$ -n-def:  $\forall n a \omega. N$ -n n a  $\omega = \text{card } \{t \in \{0..<n\}. \pi (t+1) \omega = a\}$ 
  and count-assm-pointwise:  $\forall n \omega. (\sum a \in A. \text{real } (N$ -n n a  $\omega)) = \text{real } n$ 
begin

lemma union-eq:
  fixes a :: 'a and n k :: nat
  assumes k  $\leq$  n
  shows  $\{t. t < n \wedge \pi (t+1) \omega = a\} = \{t. t < k \wedge \pi (t+1) \omega = a\} \cup \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\}$ 
proof
  show  $\{t. t < n \wedge \pi (t+1) \omega = a\} \subseteq \{t. t < k \wedge \pi (t+1) \omega = a\} \cup \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\}$ 
  proof
    fix x assume  $x \in \{t. t < n \wedge \pi (t+1) \omega = a\}$ 
    then have  $x < n$  and  $\pi (x+1) \omega = a$  by auto
    then have  $x < k \vee k \leq x$  by auto
    thus  $x \in \{t. t < k \wedge \pi (t+1) \omega = a\} \cup \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\}$ 
    using  $\langle x < n \rangle \langle \pi (x+1) \omega = a \rangle$  by auto
  qed
next
  show  $\{t. t < k \wedge \pi (t+1) \omega = a\} \cup \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\} \subseteq \{t. t < n \wedge \pi (t+1) \omega = a\}$ 
  proof
    fix x assume  $x \in \{t. t < k \wedge \pi (t+1) \omega = a\} \cup \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\}$ 
    then have  $x < k \wedge \pi (x+1) \omega = a \vee (k \leq x \wedge x < n \wedge \pi (x+1) \omega = a)$  by auto
    hence  $x < n \wedge \pi (x+1) \omega = a$  using assms by auto
    thus  $x \in \{t. t < n \wedge \pi (t+1) \omega = a\}$  by auto
  qed
qed

lemma cardinality-indic-eq:
  fixes I :: nat  $\Rightarrow$  bool
  assumes finite  $\{t. k \leq t \wedge t < n\}$ 

```

shows $\text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge I t\} = (\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0)$

proof –

have *fin*: $\text{finite } \{t. k \leq t \wedge t < n\}$ **using** *assms* **by** *simp*

have *sum-eq*:

$$\text{sum } (\text{indicator } \{t. \pi (t+1) \omega = a \wedge I t\}) \{t. k \leq t \wedge t < n\} = \text{card } (\{t. k \leq t \wedge t < n\} \cap \{t. \pi (t+1) \omega = a \wedge I t\})$$

by (*rule sum-indicator-eq-card*[*OF fin*])

have *sets-eq*:

$$\{t. k \leq t \wedge t < n\} \cap \{t. \pi (t+1) \omega = a \wedge I t\} = \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge I t\}$$

by *auto*

hence *card-eq*:

$$\text{card } (\{t. k \leq t \wedge t < n\} \cap \{t. \pi (t+1) \omega = a \wedge I t\}) = \text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge I t\}$$

by *simp*

have *card-sum-eq*:

$$\text{sum } (\text{indicator } \{t. \pi (t+1) \omega = a \wedge I t\}) \{t. k \leq t \wedge t < n\} = \text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge I t\}$$

using *sum-eq sets-eq* **by** *simp*

have *ind-eq*:

$$\text{sum } (\text{indicator } \{t. \pi (t+1) \omega = a \wedge I t\}) \{t. k \leq t \wedge t < n\} = \text{sum } (\lambda t. \text{of-bool } (\pi (t+1) \omega = a \wedge I t)) \{t. k \leq t \wedge t < n\}$$

by (*simp add: indicator-def*)

have *bool-if-eq*:

$$\text{sum } (\lambda t. \text{of-bool } (\pi (t+1) \omega = a \wedge I t)) \{t. k \leq t \wedge t < n\} = \text{sum } (\lambda t. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0) \{t. k \leq t \wedge t < n\}$$

by (*simp add: of-bool-def*)

have *set-eq*: $\{t. k \leq t \wedge t < n\} = \{k..\} \cap \{..<n\}$ **by** *auto*

have *sum-set-eq*:

$$\text{sum } (\lambda t. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0) \{t. k \leq t \wedge t < n\} = \text{sum } (\lambda t. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0) (\{k..\} \cap \{..<n\})$$

by (*simp add: set-eq*)

have *atLeastLessThan-eq*:

$$\text{sum } (\lambda t. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0) (\{k..\} \cap \{..<n\}) = \text{sum } (\lambda t. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0) (\text{atLeastLessThan } k \ n)$$

by (*simp add: atLeastLessThan-def*)

have *final-sum-eq*:

$$\text{sum } (\lambda t. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0) (\text{atLeastLessThan } k \ n) = (\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge I t \text{ then } 1 \text{ else } 0)$$

by *simp*

show *?thesis*

apply (*subst card-sum-eq*[*symmetric*])

apply (*subst ind-eq*)

apply (*subst bool-if-eq*)

apply (*subst sum-set-eq*)

apply (*subst atLeastLessThan-eq*)

apply (*simp add: final-sum-eq*)

done

qed

lemma *ge-rewrite*: $(x::real) \geq y \implies y \leq x$ by *simp*

lemma *Nn-expression*:

fixes $a :: 'a$ and $s :: nat \Rightarrow real$

and $k :: nat$ and $n :: nat$

assumes $a \in A$

and $k \leq n$

and $0 < n$

and $\forall t \in \{0..n\}. 0 < s\ t$

and $\forall t < n - 1. s\ t \leq s\ (t + 1)$

and *init-play-once*: $\forall \omega. a \in A \longrightarrow N\text{-}n\ k\ a\ \omega = 1$

and *finite-played-sets*:

finite $\{t. t < n \wedge \pi\ (t+1)\ \omega = a\}$

finite $\{t. t < k \wedge \pi\ (t+1)\ \omega = a\}$

finite $\{t. k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a\}$

shows

$(N\text{-}n\ n\ a\ \omega) = 1 + (\sum\ t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge \text{real}\ (N\text{-}n\ t\ a\ \omega) < s\ t$
then 1 else 0) +

$(\sum\ t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge \text{real}\ (N\text{-}n\ t\ a\ \omega) \geq s\ t$

else 0)

proof -

have *Nn-def*: $(N\text{-}n\ n\ a\ \omega) = \text{card}\ \{t. t < n \wedge \pi\ (t+1)\ \omega = a\}$

by (*simp add: Nn-def*)

have *init-count*: $1 \leq \text{card}\ \{t. t < k \wedge \pi\ (t+1)\ \omega = a\}$

proof -

have *eq-card*: $N\text{-}n\ k\ a\ \omega = \text{card}\ \{t. t < k \wedge \pi\ (t+1)\ \omega = a\}$

by (*simp add: Nn-def*)

moreover from *init-play-once* and $\langle a \in A \rangle$ have $1 \leq N\text{-}n\ k\ a\ \omega$

by *simp*

ultimately show *?thesis* by *simp*

qed

have *disj*: $\{t. t < k \wedge \pi\ (t+1)\ \omega = a\} \cap \{t. k \leq t \wedge t < n \wedge \pi\ (t+1)\ \omega = a\}$

= $\{\}$

by *auto*

have *union-eq-set*:

$\{t. t < n \wedge \pi\ (t+1)\ \omega = a\} = \{t. t < k \wedge \pi\ (t+1)\ \omega = a\} \cup \{t. k \leq t \wedge t <$
 $n \wedge \pi\ (t+1)\ \omega = a\}$

using *union-eq assms* by *blast*

have *finite1*: *finite* $\{t. t < k \wedge \pi\ (t+1)\ \omega = a\}$ and *finite2*: *finite* $\{t. k \leq t \wedge t <$
 $n \wedge \pi\ (t+1)\ \omega = a\}$

using *finite-played-sets* by *auto*

have *card-eq*:

$\text{card } \{t. t < n \wedge \pi (t+1) \omega = a\} = \text{card } \{t. t < k \wedge \pi (t+1) \omega = a\} + \text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\}$

using *card-Un-disjoint*[*OF finite1 finite2 disj*] *union-eq-set* **by** *simp*

have *eq-card-k*: $\text{card } \{t. t < k \wedge \pi (t+1) \omega = a\} = N\text{-}n \ k \ a \ \omega$
by (*simp add: N-n-def*)

have *card-k-eq-1*: $\text{card } \{t. t < k \wedge \pi (t+1) \omega = a\} = 1$
using *init-play-once assms eq-card-k* **by** *simp*

have *card-eq-1*:
 $\text{card } \{t. t < n \wedge \pi (t+1) \omega = a\} = 1 + \text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\}$
using *card-eq eq-card-k card-k-eq-1* **by** *simp*

have *finite-lt*: $\text{finite } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) < s \ t\}$
using *finite2 finite-subset* **by** *simp*

have *finite-ge*: $\text{finite } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) \geq s \ t\}$
using *finite2 finite-subset* **by** *simp*

have *card-eq-partition*:

$\text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\} =$
 $\text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) < s \ t\} +$
 $\text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) \geq s \ t\}$

proof –

have *union-eq-partition*:

$\{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\} =$
 $\{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) < s \ t\} \cup$
 $\{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) \geq s \ t\}$ **by** *auto*

have *disj-partition*:

$\{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) < s \ t\} \cap$
 $\{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) \geq s \ t\} = \{\}$
by *auto*

show *?thesis*

using *card-Un-disjoint*[*OF finite-lt finite-ge disj-partition*] *union-eq-partition*
by *simp*
qed

have *card-eq-sum-lt*:

$\text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) < s \ t\} =$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N\text{-}n \ t \ a \ \omega) < s \ t \text{ then } 1 \text{ else } 0)$
using *cardinality-indic-eq* **by** *simp*

have *card-eq-sum-ge*:

$\text{card } \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) \geq s \ t\} =$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) \geq s \ t \text{ then } 1 \text{ else } 0)$
using cardinality-indic-eq by simp

have $(N-n \ n \ a \ \omega) = \text{card } \{t. t < n \wedge \pi (t+1) \omega = a\}$
using N-n-def by simp

also have $\dots = 1 +$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) < s \ t \text{ then } 1 \text{ else } 0) +$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) \geq s \ t \text{ then } 1 \text{ else } 0)$
using card-eq-partition card-eq-1 card-eq-sum-lt card-eq-sum-ge by simp

also have $\dots = 1 +$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) < s \ t \text{ then } 1 \text{ else } 0) +$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge s \ t \leq \text{real } (N-n \ t \ a \ \omega) \text{ then } 1 \text{ else } 0)$
using ge-rewrite by simp

have *rewrite-eq*: $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) \geq s \ t \text{ then } 1 \text{ else } 0) =$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge s \ t \leq \text{real } (N-n \ t \ a \ \omega) \text{ then } 1 \text{ else } 0)$
using sum.cong by simp

ultimately have *final-eq*: $(N-n \ n \ a \ \omega) = 1 + (\sum t = k..<n. \text{if } \pi (t+1) \omega = a$
 $\wedge \text{real } (N-n \ t \ a \ \omega) < s \ t \text{ then } 1 \text{ else } 0) +$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) \geq s \ t \text{ then } 1$
 $\text{else } 0)$
by simp

show $(N-n \ n \ a \ \omega) = 1 + (\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega)$
 $< s \ t \text{ then } 1 \text{ else } 0) +$
 $(\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge \text{real } (N-n \ t \ a \ \omega) \geq s \ t \text{ then } 1$
 $\text{else } 0)$
apply (*subst rewrite-eq*)
apply (*subst final-eq*)
apply *simp*
done

qed

lemma *upper-bound-expression-contradiction*:

fixes $a :: 'a$ **and** $s :: \text{nat} \Rightarrow \text{real}$
and $k :: \text{nat}$ **and** $n :: \text{nat}$
and $s\text{-n-nat} :: \text{nat}$
assumes $a \in A$
and $k \leq n$
and $0 < n$
and *non-neg-s*: $\forall t \in \{0..n\}. 0 < s \ t$
and *base-le*: $s \ 0 \leq s \ 1$
and *non-dec*: $\forall t < n - 1. s \ t \leq s \ (t + 1)$

and *s-mono*: $\bigwedge t. k \leq t \wedge t \leq n \implies s t \leq s n$
and *init-play-once*: $\forall \omega. a \in A \longrightarrow N\text{-}n\ k\ a\ \omega = 1$
and *finite-played-sets*:
finite $\{t. t < n \wedge \pi (t+1)\ \omega = a\}$
finite $\{t. t < k \wedge \pi (t+1)\ \omega = a\}$
finite $\{t. k \leq t \wedge t < n \wedge \pi (t+1)\ \omega = a\}$
and *xs-sorted-def*: $xs = \text{sorted-list-of-set } \{t \in \{k..\<n\}. \pi (t+1)\ \omega = a \wedge \text{real } (N\text{-}n\ t\ a\ \omega) < s\ t\}$
and *s-nat-def*: $s\text{-}n\text{-}nat = \text{nat } (\lfloor s\ n \rfloor)$
and *len-bound-def*: $s\text{-}n\text{-}nat < \text{length } xs$
and *distinct-xs*: *distinct* xs
and *gt-ineq*: $\text{length } xs + 1 > \lfloor s\ n \rfloor$
and *N-n-increasing-with-plays*:
 $\forall t\ t'. k \leq t \wedge t < t' \wedge \pi (t+1)\ \omega = a \wedge \pi (t'+1)\ \omega = a \longrightarrow N\text{-}n\ t'\ a\ \omega \geq N\text{-}n\ t\ a\ \omega + 1$
and *neg*: $1 + \text{real } (\sum_{t=k..\<n. \text{if } \pi (t+1)\ \omega = a \wedge \text{real } (N\text{-}n\ t\ a\ \omega) < s\ t \text{ then } 1 \text{ else } 0}) > s\ n$
and *t-hat* $\in \text{set } xs$
and *t-hat* $= xs !\ s\text{-}n\text{-}nat$
and $\text{real } (N\text{-}n\ t\text{-}hat\ a\ \omega) \geq \text{real } s\text{-}n\text{-}nat + 1$

shows $(\text{real } (N\text{-}n\ t\text{-}hat\ a\ \omega) \geq \lfloor s\ n \rfloor + 1) \wedge (\pi (t\text{-}hat+1)\ \omega = a \wedge (\text{real } (N\text{-}n\ t\text{-}hat\ a\ \omega) < s\ t\text{-}hat))$

proof –

have *sn-nat-def*: $s\text{-}n\text{-}nat = \text{nat } (\lfloor s\ n \rfloor)$ **by** *fact*

let *?I* $= \{t \in \{k..\<n\}. \pi (t+1)\ \omega = a \wedge \text{real } (N\text{-}n\ t\ a\ \omega) < s\ t\}$

have *sum-eq-card*: $\text{real } (\sum_{t=k..\<n. \text{if } \pi (t+1)\ \omega = a \wedge \text{real } (N\text{-}n\ t\ a\ \omega) < s\ t \text{ then } 1 \text{ else } 0}) = \text{real } (\text{card } ?I)$

using *cardinality-indic-eq* **by** *simp*

have *finI*: *finite* $\{t \in \{k..\<n\}. \pi (t+1)\ \omega = a \wedge \text{real } (N\text{-}n\ t\ a\ \omega) < s\ t\}$

proof –

have *subset*: $\{t \in \{k..\<n\}. \pi (t+1)\ \omega = a \wedge \text{real } (N\text{-}n\ t\ a\ \omega) < s\ t\} \subseteq \{t. k \leq t \wedge t < n \wedge \pi (t+1)\ \omega = a\}$

by *auto*

moreover **have** *finite* $\{t. k \leq t \wedge t < n \wedge \pi (t+1)\ \omega = a\}$

using *finite-played-sets*(*?*) .

ultimately **show** *?thesis*

using *finite-subset* **by** *blast*

qed

have *xs-props*: *sorted* xs *distinct* xs *set* $xs = \{t \in \{k..\<n\}. \pi (t+1)\ \omega = a \wedge \text{real } (N\text{-}n\ t\ a\ \omega) < s\ t\}$

using *List.linorder-class.set-sorted-list-of-set*

using *finite-played-sets*
unfolding *xs-sorted-def*
by *auto*

have *sum-eq-length*:

$(\sum_{t = k..<n.} \text{if } \pi(t+1) \omega = a \wedge \text{real}(N-n \ t \ a \ \omega) < s \ t \ \text{then } 1 \ \text{else } 0) = \text{length } xs$

proof –

let $?P = \lambda t. \pi(t+1) \omega = a \wedge \text{real}(N-n \ t \ a \ \omega) < s \ t$

let $?I = \{t \in \{k..<n\}. ?P \ t\}$

have *sum-card*: $(\sum_{t = k..<n.} \text{if } ?P \ t \ \text{then } 1 \ \text{else } 0) = \text{card } ?I$

using *cardinality-indic-eq* **by** *simp*

also have $\dots = \text{card}(\text{set } xs)$

using *xs-props(3)* **by** *simp*

also have $\dots = \text{length } xs$

using *xs-props(2)* *distinct-card* **by** *auto*

finally show *?thesis* .

qed

have *gt-ineq*: $1 + \text{real}(\text{length } xs) > s \ n$

proof –

have *eq1*: $1 + (\sum_{t = k..<n.} \text{if } \pi(t+1) \omega = a \wedge \text{real}(N-n \ t \ a \ \omega) < s \ t \ \text{then } 1 \ \text{else } 0) = 1 + \text{length } xs$

using *sum-eq-length* **by** *simp*

then have *eq2*: $\text{real}(1 + (\sum_{t = k..<n.} \text{if } \pi(t+1) \omega = a \wedge \text{real}(N-n \ t \ a \ \omega) < s \ t \ \text{then } 1 \ \text{else } 0)) = 1 + \text{real}(\text{length } xs)$

by *simp*

from *neg* **have** $\neg(1 + \text{real}(\sum_{t=k..<n.} \text{if } \pi(t+1) \omega = a \wedge \text{real}(N-n \ t \ a \ \omega) < s \ t \ \text{then } 1 \ \text{else } 0) \leq s \ n)$

by *simp*

hence *gt*: $1 + \text{real}(\sum_{t=k..<n.} \text{if } \pi(t+1) \omega = a \wedge \text{real}(N-n \ t \ a \ \omega) < s \ t \ \text{then } 1 \ \text{else } 0) > s \ n$

by *simp*

then show *?thesis*

using *eq1 eq2 neg gt* **by** *simp*

qed

then have *len-bound-assm*: $s - n - \text{nat} < \text{length } xs$
using *len-bound-def* *assms*
by *auto*

then have *distinct-xs*: *distinct* xs
using *assms* **by** *simp*

then have *gt-ineq*: $\text{length } xs + 1 > \lfloor s \rfloor$
using *assms* **by** *simp*

have *part-1-glob*: $\pi (t\text{-hat} + 1) \omega = a \wedge \text{real } (N - n - t\text{-hat } a \omega) < s - t\text{-hat}$
proof –

have *floor-le-sn*: $\lfloor s \rfloor \leq s - n$
using *Archimedean-Field.of-int-floor-le* **by** *simp*

from *gt-ineq* **have** *len-bound*: $\text{length } xs + 1 > \lfloor s \rfloor$
using *len-bound-assm* **by** *simp*

have *nat-expression-sn*: $\text{length } xs \geq \text{nat } (\lfloor s \rfloor)$
proof –

from *len-bound* **have** $\text{length } xs \geq \text{nat } (\lfloor s \rfloor)$
by (*simp add: nat-less-iff*)

then show *?thesis* .

qed

have *sn-lt-len-plus-1*: $\lfloor s \rfloor < 1 + \text{real } (\text{length } xs)$
using *gt-ineq* **by** *simp*

then have *final-eq-glob*: $\pi (t\text{-hat} + 1) \omega = a \wedge \text{real } (N - n - t\text{-hat } a \omega) < s - t\text{-hat}$
proof –

have *pi-eq*: $\pi (t\text{-hat} + 1) \omega = a$
using *assms* *xs-sorted-def* **by** *simp*

then have $\pi (t\text{-hat} + 1) \omega = a$ **using** *pi-eq* **by** *simp*

then have *intermid-eq*: $k \leq t\text{-hat} \wedge t\text{-hat} < n \wedge \pi (t\text{-hat} + 1) \omega = a \wedge \text{real } (N - n - t\text{-hat } a \omega) < s - t\text{-hat}$
using *xs-sorted-def* $\langle t\text{-hat} \in \text{set } xs \rangle$ **by** *auto*

have *fin-eq*: $\pi (t\text{-hat} + 1) \omega = a \wedge \text{real } (N - n - t\text{-hat } a \omega) < s - t\text{-hat}$
using *intermid-eq*
by *simp*

then show *?thesis* **by** *simp*

qed

show *?thesis* **using** *final-eq-glob* **by** *simp*

qed

have *floor-le-sn*: $\lfloor s \ n \rfloor \leq s \ n$
using *Archimedean-Field.of-int-floor-le* by *simp*

have *one-plus-length-gt-floor*: $1 + \text{real}(\text{length } xs) > s \ n$
using $\langle 1 + \text{real}(\text{length } xs) > s \ n \rangle$.

have *floor-less-length-plus-one*: $\text{real}(\text{nat}(\lfloor s \ n \rfloor)) < 1 + \text{real}(\text{length } xs)$
using *one-plus-length-gt-floor floor-le-sn*
by *linarith*

from *gt-ineq* have *len-result*: $\text{length } xs > \lfloor s \ n \rfloor - 1$ by *simp*
then have $\text{length } xs \geq \text{nat}(\lfloor s \ n \rfloor)$ by (*simp add: nat-less-iff*)
hence *t-hat-in-set*: $t\text{-hat} \in \text{set } xs$
using *assms distinct-xs len-bound-assm* by *auto*

have *t-hat-le-n*: $t\text{-hat} + 1 \leq n$
using *xs-sorted-def t-hat-in-set* by *auto*

have *k-le-t-hat*: $k \leq t\text{-hat}$
using *xs-sorted-def t-hat-in-set* by *auto*

have *s-t-hat-le-s-n*: $s \ t\text{-hat} \leq s \ n$
using *s-mono[of t-hat]*
using *xs-sorted-def t-hat-in-set* by *auto*

have $\text{real}(N-n \ t\text{-hat} \ a \ \omega) \geq \text{real } s\text{-n-nat} + 1$
using *assms* by *simp*

have *final-result*: $\text{real}(N-n \ t\text{-hat} \ a \ \omega) \geq \lfloor s \ n \rfloor + 1$
using $\langle \text{real}(N-n \ t\text{-hat} \ a \ \omega) \geq \text{real}(s\text{-n-nat}) + 1 \rangle$
by (*simp add: sn-nat-def*)

show *?thesis* using *final-result* and *part-1-glob* by *auto*

qed

lemma *Nn-upper-bound*:

fixes $a :: 'a$ **and** $s :: nat \Rightarrow real$
and $k :: nat$ **and** $n :: nat$
assumes *asm*: $real(1 + (\sum t = k..<n. if \pi (t+1) \omega = a \wedge real (N-n t a \omega) < s t then 1 else 0)) \leq s n$
and *a-in-A*: $a \in A$
and *k-le-n*: $k \leq n$
and *n-pos*: $0 < n$
and *s-pos*: $\forall t \in \{0..n\}. 0 < s t$
and *s-nondec*: $\forall t < n - 1. s t \leq s (t + 1)$
and *init-play-once*: $\forall \omega. a \in A \longrightarrow N-n k a \omega = 1$
and *finite-played-sets-1*: $finite \{t. t < n \wedge \pi (t+1) \omega = a\}$
and *finite-played-sets-2*: $finite \{t. t < k \wedge \pi (t+1) \omega = a\}$
and *finite-played-sets-3*: $finite \{t. k \leq t \wedge t < n \wedge \pi (t+1) \omega = a\}$
shows $real(N-n n a \omega) \leq s n + real((\sum t = k..<n. if \pi (t+1) \omega = a \wedge s t \leq real (N-n t a \omega) then 1 else 0))$

proof –

have *expr-nat*: $N-n n a \omega =$
 $1 + (\sum t = k..<n. if \pi (t+1) \omega = a \wedge real (N-n t a \omega) < s t then 1 else 0) +$
 $(\sum t = k..<n. if \pi (t+1) \omega = a \wedge s t \leq real (N-n t a \omega) then 1 else 0)$
using *Nn-expression*[*OF a-in-A k-le-n n-pos s-pos s-nondec init-play-once*
finite-played-sets-1 finite-played-sets-2 finite-played-sets-3]
by *simp*

from *asm* **have** *bound*: $real(1 + (\sum t = k..<n. if \pi (t+1) \omega = a \wedge real (N-n t a \omega) < s t then 1 else 0)) \leq s n$ **by** *simp*

have $real (N-n n a \omega) = real (1 + (\sum t = k..<n. if \pi (t+1) \omega = a \wedge real (N-n t a \omega) < s t then 1 else 0) +$
 $(\sum t = k..<n. if \pi (t+1) \omega = a \wedge s t \leq real (N-n t a \omega) then 1 else 0))$
using *expr-nat* **by** *simp*

also have $\dots \leq s n + real ((\sum t = k..<n. if \pi (t+1) \omega = a \wedge s t \leq real (N-n t a \omega) then 1 else 0))$
using *assms*
by *simp*

then show *?thesis*
using *expr-nat asm*
by *simp*

qed

theorem *ENn-upper-bound*:

assumes
a-in-A: $a \in A$
and *k-le-n*: $k \leq n$
and *n-pos*: $0 < n$
and *s-pos*: $\forall t \in \{0..n\}. 0 < s t$

and *s-nondec*: $\forall t < n. s\ t \leq s\ (t + 1)$
and *init-play-once*: $\forall \omega. a \in A \longrightarrow N\text{-}n\ k\ a\ \omega = 1$
and *integrable-Nn*: *integrable* $M\ (\lambda\omega. \text{real}\ (N\text{-}n\ n\ a\ \omega))$
and *integrable-rhs-sum*: *integrable* $M\ (\lambda\omega. s\ n + (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0))$
and *integrable-s*: *integrable* $M\ (\lambda\omega. s\ n)$
and *integrable-indicator-sum*: *integrable* $M\ (\lambda\omega. \sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0)$
and *linearity*: *integral*^L $M\ (\lambda\omega. s\ n + (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0)) =$
integral^L $M\ (\lambda\omega. s\ n) + \text{integral}^L\ M\ (\lambda\omega. \sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0)$
and *pointwise-bound*: $\text{real}\ (N\text{-}n\ n\ a\ \omega) \leq s\ n + (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0)$
and *mono-intgrl*: *integral*^L $M\ (\lambda\omega. \text{real}\ (N\text{-}n\ n\ a\ \omega)) \leq \text{integral}^L\ M\ (\lambda\omega. s\ n + (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0))$
shows
expectation $(\lambda\omega. \text{real}\ (N\text{-}n\ n\ a\ \omega)) \leq s\ n + \text{expectation}\ (\lambda\omega. (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0))$
proof –
let $?f = \lambda\omega. \text{real}\ (N\text{-}n\ n\ a\ \omega)$
let $?g = \lambda\omega. s\ n + (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0)$
let $?g1 = \lambda\omega. s\ n$
let $?g2 = \lambda\omega. (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0)$
have *pointwise-le*: $?f\ \omega \leq ?g\ \omega$
using *pointwise-bound* .

have *intg-f*: *integrable* $M\ ?f$ **using** *integrable-Nn* .
have *intg-g*: *integrable* $M\ ?g$ **using** *integrable-rhs-sum* .
then **have** *eq-f*: *integral*^L $M\ ?f = \text{expectation}\ (?f)$
using *prob-space* **by** *simp*
then **have** *eq-g*: *integral*^L $M\ ?g = \text{expectation}\ (?g)$
by *simp*

then **have** *sub-egq1*: *integral*^L $M\ ?g1 = \text{expectation}\ (\lambda\omega. s\ n)$
by *simp*
then **have** *sub-egq2*: *integral*^L $M\ ?g2 = \text{expectation}\ (\lambda\omega. (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0))$
by *simp*

have $\text{real}\ (N\text{-}n\ n\ a\ \omega) \leq s\ n + (\sum t = k..<n. \text{if}\ \pi\ (t+1)\ \omega = a \wedge s\ t \leq \text{real}\ (N\text{-}n\ t\ a\ \omega)\ \text{then}\ 1\ \text{else}\ 0)$
using *assms* **by** *simp*

```

have const-measure: Sigma-Algebra.measure M (space M) = 1
  using prob-space Probability-Measure.prob-space.prob-space by blast

have exp-sn: expectation (λω. s n) = s n
proof -
  have expectation (λω. s n) = integralL M (λω. s n)
    by simp
  also have ... = s n
    using prob-space by (simp add: prob-space.prob-space)
  then show ?thesis .
qed

have rhs-expr:
  integralL M ?g = integralL M (λω. s n) + integralL M (λω. ∑ t = k.. $n$ . if
π (t+1) ω = a ∧ s t ≤ real (N-n t a ω) then 1 else 0)
  using linearity by auto

then have rhs-expr:
  expectation (?g) = expectation (λω. s n) + expectation (λω. (∑ t = k.. $n$ . if
π (t+1) ω = a ∧ s t ≤ real (N-n t a ω) then 1 else 0))
  by simp

also have final-eq: expectation (?g) = s n + expectation (λω. (∑ t = k.. $n$ . if
π (t+1) ω = a ∧ s t ≤ real (N-n t a ω) then 1 else 0))
  using rhs-expr exp-sn
  by simp

have glob-final-eq: expectation (?f) ≤ expectation (?g)
  using asms intg-f intg-g pointwise-le mono-intgrl
  by auto

thus ?thesis
  using glob-final-eq final-eq by simp
qed

end
end
theory Discrete-UCB-Step2
  imports Discrete-UCB-Step1

begin

locale bandit-policy = bandit-problem + prob-space +
  fixes Ω :: 'b set
    and ℱ :: 'b set set
    and ω :: 'b
  fixes π :: nat ⇒ 'b ⇒ 'a
    and N-n :: nat ⇒ 'a ⇒ 'b ⇒ nat

```

```

and  $Z :: \text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$ 
and  $\delta :: \text{real}$ 
and  $q :: \text{real}$ 
assumes finite-A:  $\text{finite } A$ 
and a-in-A:  $a \in A$ 
and measurable-policy:  $\forall t. \pi t \in \text{measurable } M \text{ (count-space } A)$ 
and N-n-def:  $\forall n a b. N\text{-n } n a b = \text{card } \{t \in \{0..<n\}. \pi (t+1) b = a\}$ 
and delta-pos:  $0 < \delta$ 
and delta-less1:  $\delta < 1$ 
and q-pos:  $q > 0$ 

begin

definition sample-mean-Z ::  $\text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$  where
  sample-mean-Z  $n a b \equiv (1 / \text{real } n) * (\sum_{i < n}. Z i a b)$ 

definition M-val ::  $\text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$  where
  M-val  $t a b \equiv (\text{if } N\text{-n } (t+1) a b = 0 \text{ then } 0$ 
     $\text{else } (\sum_{s < t}. \text{if } \pi s b = a \text{ then } Z s a b \text{ else } 0) / \text{real } (N\text{-n } t a b))$ 

definition U ::  $\text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$  where
  U  $t a b \equiv M\text{-val } t a b + q * \text{sqrt } (\ln (1 / \delta) / (2 * \text{real } (\max 1 (N\text{-n } t a b))))$ 

definition A-t-plus-1 ::  $\text{nat} \Rightarrow 'b \Rightarrow 'a$  where
  A-t-plus-1  $t b \equiv (\text{SOME } a. a \in A \wedge (\forall a'. a' \in A \longrightarrow U t a b \geq U t a' b))$ 

lemma (in finite-measure) finite-measure-mono:
  assumes  $A \subseteq B$   $B \in \text{sets } M$  shows  $\text{measure } M A \leq \text{measure } M B$ 
  using emeasure-mono[OF assms] emeasure-real[of A] emeasure-real[of B] by
  (auto simp: measure-def)

theorem union-bound:
  fixes  $E F G :: 'b \text{ set}$ 
  assumes  $E \subseteq F \cup G$ 
  and  $E \in \text{events } F \in \text{events } G \in \text{events}$ 
  shows  $\text{prob } E \leq \text{prob } F + \text{prob } G$ 
proof –
  have  $F \cup G \in \text{events}$ 
  using assms(3,4) sets.Un by blast
  have  $\text{prob } E \leq \text{prob } (F \cup G)$ 
  using assms local.finite-measure-mono by auto
  also have  $\text{prob } (F \cup G) \leq \text{prob } F + \text{prob } G$ 
  using assms measure-Un-le by blast
  finally show ?thesis .
qed

```

theorem *hoeffding-iid-bound-ge-general*:
fixes $a :: 'a$ **and** $n :: nat$ **and** $\varepsilon :: real$ **and** $\mu\text{-hat} :: real$ **and** $l\ u :: real$
assumes *a-in*: $a \in A$
and *eps-pos*: $\varepsilon \geq 0$
and *bounds*: $\forall i < n. \forall \omega \in \Omega. l \leq Z\ i\ a\ \omega \wedge Z\ i\ a\ \omega \leq u$
and *mu-def*: $\mu\text{-hat} = (\sum\ i < n. expectation\ (\lambda\omega. Z\ i\ a\ \omega))$
and $u - l \neq 0$
and *n-pos*: $n > 0$
and *space-M*: $space\ M = \Omega$
and *sets-M*: $sets\ M = \mathcal{F}$
and *indep*: *indep-vars* $(\lambda\cdot. borel)\ (\lambda i. (\lambda\omega. Z\ i\ a\ \omega))\ \{i. i < n\}$
and *rv*: $\forall i < n. random\ variable\ borel\ (\lambda\omega. Z\ i\ a\ \omega)$

shows $prob\ \{\omega \in \Omega. (\sum\ i < n. Z\ i\ a\ \omega) \geq \mu\text{-hat} + \varepsilon\}$
 $\leq exp\ (-2 * \varepsilon^2 / (real\ n * (u - l)^2))$

proof –

let $?I = \{i. i < n\}$
let $?X = \lambda i. (\lambda\omega. Z\ i\ a\ \omega)$
let $?a = \lambda i. l$
let $?b = \lambda i. u$

have *finite-I*: *finite* $?I$ **by** *simp*

have *AE-bounds*: $\forall i \in ?I. AE\ \omega\ in\ M. ?a\ i \leq ?X\ i\ \omega \wedge ?X\ i\ \omega \leq ?b\ i$

proof

fix i **assume** $iI: i \in ?I$
have $\forall \omega \in \Omega. l \leq Z\ i\ a\ \omega \wedge Z\ i\ a\ \omega \leq u$ **using** *bounds* iI **by** *simp*
thus $AE\ \omega\ in\ M. ?a\ i \leq ?X\ i\ \omega \wedge ?X\ i\ \omega \leq ?b\ i$
using *assms* **by** *auto*

qed

have *indep-loc*: *indep-interval-bounded-random-variables* $M\ ?I\ ?X\ ?a\ ?b$
by (*standard*; *use* *finite-I* *indep* *AE-bounds* **in** *auto*)

from *indep-loc*

have H : *Hoeffding-ineq* $M\ ?I\ ?X\ ?a\ ?b$
by (*rule* *Hoeffding-ineq.intro*)

have *widths*: $(\sum\ i \in ?I. (?b\ i - ?a\ i)^2) = real\ n * (u - l)^2$
by *simp*

have *widths-pos*: $0 < (\sum\ i \in ?I. (?b\ i - ?a\ i)^2)$
using $\langle u - l \neq 0 \rangle$ *n-pos* **by** *simp*

have *sum-expectations-eq-integrals*:

$(\sum\ i < n. expectation\ (\lambda\omega. Z\ i\ a\ \omega)) = (\sum\ i \in ?I. integral^L\ M\ (?X\ i))$

proof –

have *eq1*: $\forall i < n. expectation\ (\lambda\omega. Z\ i\ a\ \omega) = integral^L\ M\ (?X\ i)$

using *rv space-M sets-M* **by** *simp*
moreover have *sum-eq*: $(\sum_{i < n}. \text{integral}^L M (?X i)) = (\sum_{i \in ?I}. \text{integral}^L M (?X i))$
by (*simp add: lessThan-def*)
ultimately show *?thesis*
by *simp*
qed

have *sum-integrals-eq*: $(\sum_{i \in ?I}. \text{integral}^L M (?X i)) = \mu\text{-hat}$
using *mu-def* **by** (*simp add: sum-expectations-eq-integrals*)

have *tail*:
 $\text{prob } \{\omega \in \Omega. (\sum_{i \in ?I}. ?X i \omega) - (\sum_{i \in ?I}. \text{integral}^L M (?X i)) \geq \varepsilon\}$
 $\leq \exp(-2 * \varepsilon^2 / (\sum_{i \in ?I}. (?b i - ?a i)^2))$
using *Hoeffding-ineq.Hoeffding-ineq-ge[OF H eps-pos widths-pos]*
by (*smt (verit, best) Collect-cong assms(7)*)

have *lhs-rewrite*: $\{\omega \in \Omega. (\sum_{i < n}. Z i a \omega) \geq \mu\text{-hat} + \varepsilon\}$
 $= \{\omega \in \Omega. (\sum_{i \in ?I}. ?X i \omega) - (\sum_{i \in ?I}. \text{integral}^L M (?X i)) \geq \varepsilon\}$
by (*simp add: add commute le-diff-eq lessThan-def sum-integrals-eq*)

have *rhs-rewrite*: $\exp(-2 * \varepsilon^2 / (\sum_{i \in ?I}. (?b i - ?a i)^2))$
 $= \exp(-2 * \varepsilon^2 / (\text{real } n * (u - l)^2))$
using *widths* **by** *simp*

show *?thesis*
using *tail lhs-rewrite rhs-rewrite* **by** *simp*
qed

theorem *hoeffding-iid-bound-le-general*:

fixes *a* :: 'a **and** *n* :: nat **and** ε :: real **and** $\mu\text{-hat}$:: real **and** *l u* :: real
assumes *a-in*: $a \in A$

and *eps-pos*: $\varepsilon \geq 0$

and *bounds*: $\forall i < n. \forall \omega \in \Omega. l \leq Z i a \omega \wedge Z i a \omega \leq u$

and *mu-def*: $\mu\text{-hat} = (\sum_{i < n}. \text{expectation } (\lambda \omega. Z i a \omega))$

and $u - l \neq 0$

and *n-pos*: $n > 0$

and *space-M*: $\text{space } M = \Omega$

and *sets-M*: $\text{sets } M = \mathcal{F}$

and *indep*: *indep-vars* ($\lambda \cdot. \text{borel}$) ($\lambda i. (\lambda \omega. Z i a \omega)$) $\{i. i < n\}$

and *rv*: $\forall i < n. \text{random-variable borel } (\lambda \omega. Z i a \omega)$

shows $\text{prob } \{\omega \in \Omega. (\sum_{i < n}. Z i a \omega) \leq \mu\text{-hat} - \varepsilon\}$
 $\leq \exp(-2 * \varepsilon^2 / (\text{real } n * (u - l)^2))$

proof –

let *?I* = $\{i. i < n\}$

let $?X = \lambda i. (\lambda \omega. Z i a \omega)$
let $?a = \lambda i. l$
let $?b = \lambda i. u$

have *finite-I*: *finite ?I by simp*

have *AE-bounds*: $\forall i \in ?I. AE \ \omega \text{ in } M. ?a \ i \leq ?X \ i \ \omega \wedge ?X \ i \ \omega \leq ?b \ i$
proof

fix i **assume** $iI: i \in ?I$
have $\forall \omega \in \Omega. l \leq Z \ i \ a \ \omega \wedge Z \ i \ a \ \omega \leq u$ **using** *bounds iI by simp*
thus $AE \ \omega \text{ in } M. ?a \ i \leq ?X \ i \ \omega \wedge ?X \ i \ \omega \leq ?b \ i$
using *assms by auto*

qed

have *indep-loc*: *indep-interval-bounded-random-variables M ?I ?X ?a ?b*
by (*standard; use finite-I indep AE-bounds in auto*)

from *indep-loc*

have H : *Hoeffding-ineq M ?I ?X ?a ?b*
by (*rule Hoeffding-ineq.intro*)

have *widths*: $(\sum i \in ?I. (?b \ i - ?a \ i)^2) = real \ n * (u - l)^2$ **by** *simp*
have *widths-pos*: $0 < (\sum i \in ?I. (?b \ i - ?a \ i)^2)$ **using** $\langle u - l \neq 0 \rangle$ *n-pos by simp*

have *sum-expectations-eq-integrals*:

$(\sum i < n. expectation (\lambda \omega. Z i a \omega)) = (\sum i \in ?I. integral^L M (?X \ i))$

proof –

have *eq1*: $\forall i < n. expectation (\lambda \omega. Z i a \omega) = integral^L M (?X \ i)$ **using** *rv space-M sets-M by simp*

moreover **have** *sum-eq*: $(\sum i < n. integral^L M (?X \ i)) = (\sum i \in ?I. integral^L M (?X \ i))$ **by** (*simp add: lessThan-def*)

ultimately show *?thesis by simp*

qed

have *sum-integrals-eq*: $(\sum i \in ?I. integral^L M (?X \ i)) = \mu\text{-hat}$
using *mu-def by (simp add: sum-expectations-eq-integrals)*

have *tail*:

$prob \ \{\omega \in \Omega. (\sum i \in ?I. ?X \ i \ \omega) - (\sum i \in ?I. integral^L M (?X \ i)) \leq - \ \varepsilon\}$
 $\leq exp \ (- \ 2 * \ \varepsilon^2 / (\sum i \in ?I. (?b \ i - ?a \ i)^2))$

using *Hoeffding-ineq.Hoeffding-ineq-le[OF H eps-pos widths-pos]*

by (*smt (verit, best) Collect-cong assms(7)*)

have *lhs-rewrite*: $\{\omega \in \Omega. (\sum i < n. Z i a \omega) \leq \mu\text{-hat} - \varepsilon\}$
 $= \{\omega \in \Omega. (\sum i \in ?I. ?X i \omega) - (\sum i \in ?I. \text{integral}^L M (?X i))$
 $\leq -\varepsilon\}$
using *add.inverse-inverse*[of ε] *add.inverse-inverse*[of $\mu\text{-hat}$] *assms*(4)
cancel-ab-semigroup-add-class.diff-right-commute[of $\mu\text{-hat}$ $\mu\text{-hat}$ $\sum uuc < n. Z$
 $uuc a$ -]
cancel-ab-semigroup-add-class.diff-right-commute[of $-\mu\text{-hat}$ $-\mu\text{-hat}$ $\varepsilon -$
 $\mu\text{-hat}$]
cancel-ab-semigroup-add-class.diff-right-commute[of $\varepsilon - \mu\text{-hat}$ $\varepsilon - \mu\text{-hat}$ ε]
cancel-ab-semigroup-add-class.diff-right-commute[of $\varepsilon \varepsilon \mu\text{-hat}$]
cancel-ab-semigroup-add-class.diff-right-commute[of $0 - \mu\text{-hat}$ ε]
cancel-ab-semigroup-add-class.diff-right-commute[of $\mu\text{-hat}$ $\mu\text{-hat}$ ε]
cancel-ab-semigroup-add-class.diff-right-commute[of $-\mu\text{-hat}$ $-\mu\text{-hat}$ $(\sum uuc < n.$
 $Z uuc a -) - \mu\text{-hat}$]
cancel-ab-semigroup-add-class.diff-right-commute[of $(\sum uuc < n. Z uuc a -) -$
 $\mu\text{-hat}$ $(\sum uuc < n. Z uuc a -) - \mu\text{-hat}$
 $\sum uuc < n. Z uuc a -]$
cancel-ab-semigroup-add-class.diff-right-commute[of $\sum uuc < n. Z uuc a -$
 $\sum uuc < n. Z uuc a - \mu\text{-hat}$]
cancel-ab-semigroup-add-class.diff-right-commute[of $0 - \mu\text{-hat}$ $\sum uuc < n. Z$
 $uuc a -]$
cancel-comm-monoid-add-class.diff-cancel[of $\varepsilon - \mu\text{-hat}$] *cancel-comm-monoid-add-class.diff-cancel*[of
 ε]
cancel-comm-monoid-add-class.diff-cancel[of $\mu\text{-hat}$] *cancel-comm-monoid-add-class.diff-cancel*[of
 $-\mu\text{-hat}$]
cancel-comm-monoid-add-class.diff-cancel[of $(\sum uuc < n. Z uuc a -) - \mu\text{-hat}$]
cancel-comm-monoid-add-class.diff-cancel[of $\sum uuc < n. Z uuc a -]$ *diff-0*
diff-right-mono[of $\varepsilon \mu\text{-hat} - (\sum uuc < n. Z uuc a -) \mu\text{-hat}$] *diff-right-mono*[of
 $\sum uuc < n. Z uuc a - \mu\text{-hat} - \varepsilon \mu\text{-hat}$]
lessThan-def[of n] *more-arith-simps*(1)[of $-\varepsilon (\sum uuc < n. Z uuc a -) - \mu\text{-hat}$]
by force

have *rhs-rewrite*: $\exp(-2 * \varepsilon^2 / (\sum i \in ?I. (?b i - ?a i)^2)) = \exp(-2 * \varepsilon^2 / (\text{real } n * (u - l)^2))$
using *widths by simp*

show *?thesis using tail lhs-rewrite rhs-rewrite by simp*
qed

theorem *hoeffding-iid-ge-delta-bound*:

fixes $a :: 'a$ **and** $n :: \text{nat}$ **and** $\delta\text{-hat} :: \text{real}$ **and** $\mu\text{-hat} :: \text{real}$ **and** $l u :: \text{real}$
assumes *a-in*: $a \in A$
and *delta-bound*: $0 < \delta\text{-hat}$ $\delta\text{-hat} \leq 1$
and *bounds*: $\forall i < n. \forall \omega \in \Omega. l \leq Z i a \omega \wedge Z i a \omega \leq u$
and *mu-def*: $\mu\text{-hat} = (\sum i < n. \text{expectation } (\lambda \omega. Z i a \omega))$
and *n-pos*: $n > 0$
and *eps-pos*: $\varepsilon \geq 0$
and *u-minus-l-nonzero*: $u - l \neq 0$

```

and space-M: space M = Ω
and sets-M: sets M = ℱ
and indep: indep-vars (λ-. borel) (λi. (λω. Z i a ω)) {i. i < n}
and rv: ∀ i < n. random-variable borel (λω. Z i a ω)
and eps-expression: ε = sqrt ((real n * (u - l)^2 * ln (1 / δ-hat)) / 2)
shows prob {ω ∈ Ω. (∑ i < n. Z i a ω) ≥ μ-hat + ε} ≤ δ-hat ∧
      prob {ω ∈ Ω. (∑ i < n. Z i a ω) ≤ μ-hat - ε} ≤ δ-hat
proof -
  have eps-pos: ε ≥ 0
    using assms(7) by auto
  have eps-squared: ε^2 = (real n * (u - l)^2 * ln (1 / δ-hat)) / 2
    using eps-expression by (simp add: assms(3) delta-bound(1))

  have exp-eq: exp (- 2 * ε^2 / (real n * (u - l)^2)) = δ-hat
  proof -
    have - 2 * ε^2 / (real n * (u - l)^2) = - ln (1 / δ-hat)
    proof -
      have ε^2 / (real n * (u - l)^2) = (ln (1 / δ-hat)) / 2
        using assms(6,8) eps-squared by fastforce
      then show ?thesis by linarith
    qed
  also have ... = ln δ-hat
    using delta-bound(1) by (simp add: ln-div)
  finally show ?thesis
    using delta-bound(1) by simp
  qed

have ge-bound:
  prob {ω ∈ Ω. (∑ i < n. Z i a ω) ≥ μ-hat + ε} ≤ δ-hat
  using hoeffding-iid-bound-ge-general[OF a-in eps-pos bounds mu-def u-minus-l-nonzero

    n-pos space-M sets-M indep rv] exp-eq by simp

have le-bound:
  prob {ω ∈ Ω. (∑ i < n. Z i a ω) ≤ μ-hat - ε} ≤ δ-hat
  using hoeffding-iid-bound-le-general[OF a-in eps-pos bounds mu-def u-minus-l-nonzero

    n-pos space-M sets-M indep rv] exp-eq by simp

show ?thesis
  using ge-bound le-bound by simp
qed

lemma add-le-iff:
  fixes x y z :: real
  shows x ≤ y - z ↔ x - y ≤ -z
  by auto
lemma max-Suc-0-eq-1: max (Suc 0) x = max 1 x
  by simp

```

theorem *ucb-suboptimal-bound-set*:

fixes $t :: \text{nat}$
and $a :: 'a$
and $\Delta :: 'a \Rightarrow \text{real}$
assumes *finite-A*: $\text{finite } A$
and *a-in-A*: $a \in A$
and *a-star-in-A*: $a\text{-star} \in A$
and *argmax-exists*: $A \neq \{\}$
and *subopt-gap*: $\Delta a > 0$
and *a-not-opt*: $\exists a'. a' \in A \wedge \Delta a > 0$
and *delta-a*: $\Delta a = \mu a\text{-star} - \mu a$
and *omega-in-Omega*: $\omega \in \Omega$
and *asm*: $\omega \in \{\omega \in \Omega. A\text{-t-plus-1 } t \omega = a\}$
and *setopt*: $\forall \omega \in \Omega. \exists a\text{-max} \in A. \forall b \in A. U t b \omega \leq U t a\text{-max } \omega$
and *A-t-plus-1-maximizes*:
 $\wedge t \omega a. A\text{-t-plus-1 } t \omega = a \implies a \in A \wedge (\forall b \in A. U t a \omega \geq U t b \omega)$
shows $\{\omega \in \Omega. A\text{-t-plus-1 } t \omega = a\} \subseteq$
 $\{\omega \in \Omega. U t a\text{-star } \omega \leq \mu a\text{-star}\} \cup \{\omega \in \Omega. \mu a\text{-star} \leq U t a \omega\}$
proof –
have *set-result-1*:
 $\{\omega \in \Omega. A\text{-t-plus-1 } t \omega = a\} \subseteq \{\omega \in \Omega. \forall b \in A. U t a \omega \geq U t b \omega\}$
proof
fix ω
assume $\omega \in \{\omega \in \Omega. A\text{-t-plus-1 } t \omega = a\}$
hence $A\text{-t-plus-1 } t \omega = a$ **by** *simp*

from *setopt* **obtain** *a-max* **where**
 $a\text{-max} \in A \wedge (\forall b \in A. U t b \omega \leq U t a\text{-max } \omega)$
using $\langle \omega \in \{\omega \in \Omega. A\text{-t-plus-1 } t \omega = a\} \rangle$ **by** *auto*

hence $\forall b \in A. U t a \omega \geq U t b \omega$
using $\langle A\text{-t-plus-1 } t \omega = a \rangle$ *A-t-plus-1-maximizes* **by** *auto*

thus $\omega \in \{\omega \in \Omega. \forall b \in A. U t a \omega \geq U t b \omega\}$
using $\langle \omega \in \{\omega \in \Omega. A\text{-t-plus-1 } t \omega = a\} \rangle$ **by** *blast*
qed

have *set-result-2*: $\{\omega \in \Omega. \forall b \in A. U t a \omega \geq U t b \omega\} \subseteq \{\omega \in \Omega. U t a\text{-star } \omega \leq U t a \omega\}$

proof

fix ω **assume** *asm*: $\omega \in \{\omega \in \Omega. \forall b \in A. U t a \omega \geq U t b \omega\}$
hence $\omega \in \Omega$ **and** *ub*: $\forall b \in A. U t a \omega \geq U t b \omega$ **by** *simp-all*
from *a-star-in-A* **have** $U t a \omega \geq U t a\text{-star } \omega$ **using** *ub* **by** *simp*

```

thus  $\omega \in \{\omega \in \Omega. U t a\text{-star } \omega \leq U t a \omega\}$  using  $\langle \omega \in \Omega \rangle$  by simp
qed

have set-result-3:  $\{\omega \in \Omega. U t a\text{-star } \omega \leq U t a \omega\} \subseteq$ 
 $\{\omega \in \Omega. U t a\text{-star } \omega \leq \mu a\text{-star}\} \cup \{\omega \in \Omega. \mu a\text{-star} \leq U t a \omega\}$ 
proof
  fix  $\omega$  assume asm:  $\omega \in \{\omega \in \Omega. U t a\text{-star } \omega \leq U t a \omega\}$ 
  hence  $\omega \in \Omega$  and le:  $U t a\text{-star } \omega \leq U t a \omega$  by simp-all
  show  $\omega \in \{\omega \in \Omega. U t a\text{-star } \omega \leq \mu a\text{-star}\} \cup \{\omega \in \Omega. \mu a\text{-star} \leq U t a \omega\}$ 
  proof (cases  $U t a\text{-star } \omega \leq \mu a\text{-star}$ )
    case True
      then show ?thesis using  $\langle \omega \in \Omega \rangle$  by simp
    next
      case False
        hence  $\mu a\text{-star} < U t a\text{-star } \omega$  by simp
        with le have  $\mu a\text{-star} < U t a \omega$  by (simp add: less-le-trans)
        then show ?thesis using  $\langle \omega \in \Omega \rangle$  by simp
  qed
qed

from set-result-1 set-result-2 set-result-3 show ?thesis by auto
qed

```

```

theorem ucb-suboptimal-bound-prob-statement:
  fixes  $t :: \text{nat}$  and  $a :: 'a$  and  $\Delta :: 'a \Rightarrow \text{real}$ 
  assumes finite-A: finite  $A$ 
    and a-star-in-A:  $a\text{-star} \in A$ 
    and argmax-exists:  $A \neq \{\}$ 
    and subopt-gap:  $\Delta a > 0$ 
    and a-not-opt:  $\exists a'. a' \in A \wedge \Delta a > 0$ 
    and  $\omega$ -in- $\Omega$ :  $\omega \in \Omega$ 
    and asm:  $\omega \in \{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\}$ 
    and setopt:  $\forall \omega \in \Omega. \exists a\text{-max} \in A. \forall b \in A. U t b \omega \leq U t a\text{-max } \omega$ 
    and A-t-plus-1-maximizes:
 $\bigwedge t \omega a. A\text{-}t\text{-plus-1 } t \omega = a \implies a \in A \wedge (\forall b \in A. U t a \omega \geq U t b \omega)$ 
    and a-in-A:  $a \in A$ 
    and omega-in:  $\omega \in \Omega$ 
    and subopt-gap:  $\Delta a > 0$ 
    and delta-a:  $\Delta a = \mu a\text{-star} - \mu a$ 
    and H-def:  $H = (2 * \ln (1 / \delta)) / (\Delta a)^2$ 
    and E-def:  $E = \{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n t a \omega)\}$ 
    and F-def:  $F = \{\omega \in \Omega. U t a\text{-star } \omega \leq \mu a\text{-star}\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n t a \omega)\}$ 
    and G-def:  $G = \{\omega \in \Omega. \mu a\text{-star} \leq U t a \omega\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n t a \omega)\}$ 
    and meas-sets:  $E \in \text{sets } M F \in \text{sets } M G \in \text{sets } M$ 
    and prob-inter:  $\text{prob } (F \cap G) \equiv \text{enn2real } (\text{emeasure } M (F \cap G))$ 

```

shows $\text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\}) \leq$
 $\text{prob } (\{\omega \in \Omega. U \text{ } t \text{ } a\text{-star } \omega \leq \mu \text{ } a\text{-star}\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\})$
+
 $\text{prob } (\{\omega \in \Omega. U \text{ } t \text{ } a \omega \geq \mu \text{ } a\text{-star}\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\})$
proof –

have *subset-result*: $E \subseteq F \cup G$

proof –

have *step1*: $\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \subseteq$
 $\{\omega \in \Omega. U \text{ } t \text{ } a\text{-star } \omega \leq \mu \text{ } a\text{-star}\} \cup \{\omega \in \Omega. \mu \text{ } a\text{-star} \leq U \text{ } t \text{ } a \omega\}$
using *ucb-suboptimal-bound-set* *assms* **by** *blast*

then have *step2*:

$\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\} \subseteq$
 $(\{\omega \in \Omega. U \text{ } t \text{ } a\text{-star } \omega \leq \mu \text{ } a\text{-star}\} \cup \{\omega \in \Omega. \mu \text{ } a\text{-star} \leq U \text{ } t \text{ } a \omega\}) \cap \{\omega \in$
 $\Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\}$
by *auto*

then have *step3*:

$\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\} \subseteq$
 $(\{\omega \in \Omega. U \text{ } t \text{ } a\text{-star } \omega \leq \mu \text{ } a\text{-star}\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\}) \cup$
 $(\{\omega \in \Omega. \mu \text{ } a\text{-star} \leq U \text{ } t \text{ } a \omega\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\})$
by (*auto simp add: set-eq-iff*)

then have *step4*:

$E \subseteq F \cup G$ **using** *assms* **by** *simp*

then show *?thesis*

by (*simp add: E-def F-def G-def*)

qed

have *bound*: $\text{prob } E \leq \text{prob } F + \text{prob } G$

proof (*rule union-bound*)

show $E \subseteq F \cup G$ **by** (*simp add: subset-result*)

show $E \in \text{sets } M$ **using** *meas-sets* **by** *simp*

show $F \in \text{sets } M$ **using** *meas-sets* **by** *simp*

show $G \in \text{sets } M$ **using** *meas-sets* **by** *simp*

have $\text{prob } (F \cap G) \equiv \text{enn2real } (\text{emeasure } M (F \cap G))$ **using** *prob-inter* **by**
simp

have $\text{prob } E \leq \text{prob } (F \cup G)$

using *assms(18,19,20)* *increasingD* *measure-increasing* *subset-result* **by** *blast*

have $\text{prob } (F \cup G) = \text{prob } F + \text{prob } G - \text{prob } (F \cap G)$

by (*simp add: Int-commute assms(19,20)* *finite-measure-Diff'* *finite-measure-Union'*)

qed

hence *union-bound*:

$\text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\}) \leq$
 $\text{prob } (\{\omega \in \Omega. U \text{ } t \text{ } a\text{-star } \omega \leq \mu \text{ } a\text{-star}\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\}) +$
 $\text{prob } (\{\omega \in \Omega. U \text{ } t \text{ } a \omega \geq \mu \text{ } a\text{-star}\} \cap \{\omega \in \Omega. H \leq \text{real } (N\text{-}n \text{ } t \text{ } a \omega)\})$

using *bound assms by simp*

show *?thesis*

using *union-bound by simp*

qed

lemma *U-le- μ -pointwise:*

$U\ t\ a\text{-star}\ \omega \leq \mu\ a\text{-star} \longleftrightarrow$
 $M\text{-val}\ t\ a\text{-star}\ \omega - \mu\ a\text{-star} \leq$
 $- q * \text{sqrt} (\ln (1 / \delta) / (2 * \text{real} (\max 1 (N\text{-n}\ t\ a\text{-star}\ \omega))))$

unfolding *U-def*

using *add-le-iff max-Suc-0-eq-1*

by *auto*

lemma *U-ge- μ -pointwise:*

assumes *delta-a: $\Delta\ a = \mu\ a\text{-star} - \mu\ a$*

shows

$U\ t\ a\ \omega \geq \mu\ a\text{-star} \longleftrightarrow$
 $M\text{-val}\ t\ a\ \omega - \mu\ a \geq \Delta\ a - q * \text{sqrt} (\ln (1 / \delta) / (2 * \text{real} (\max 1 (N\text{-n}\ t\ a\ \omega))))$

proof *-*

have *delta-plus-mu-eq:*

$\Delta\ a + \mu\ a = \mu\ a\text{-star}$

proof *-*

have $\Delta\ a = \mu\ a\text{-star} - \mu\ a$

using *delta-a by simp*

hence $\Delta\ a + \mu\ a = (\mu\ a\text{-star} - \mu\ a) + \mu\ a$

by *simp*

also have $\dots = \mu\ a\text{-star}$

by *simp*

finally show *?thesis .*

qed

have *U-ge- μ -rewrite:*

$U\ t\ a\ \omega \geq \mu\ a\text{-star} \longleftrightarrow M\text{-val}\ t\ a\ \omega - \mu\ a \geq \Delta\ a - q * \text{sqrt} (\ln (1 / \delta) /$
 $(2 * \text{real} (\max 1 (N\text{-n}\ t\ a\ \omega))))$

unfolding *U-def using add-le-iff max-Suc-0-eq-1 delta-plus-mu-eq by auto*

then show *?thesis by simp*

qed

theorem *hoeffding-iid-bound-le:*

fixes $a :: 'a$ **and** $n :: \text{nat}$ **and** $\varepsilon :: \text{real}$ **and** $\mu\text{-hat} :: \text{real}$ **and** $l\text{-hat}\ u\text{-hat} :: \text{real}$

and $I :: \text{nat set}$

and $X\text{-new} :: \text{nat} \Rightarrow 'b \Rightarrow \text{real}$

and $a\text{-bound}\ b\text{-bound} :: \text{nat} \Rightarrow \text{real}$

assumes $a\text{-in}: a \in A$

and $b \in \Omega$
and $eps\text{-pos}: \varepsilon \geq 0$
and $eps: \varepsilon = \text{abs } (u\text{-hat} - l\text{-hat}) * \text{sqrt } (((\text{real } n) / 2) * \ln (1 / \delta))$
and $\delta \geq 0 \wedge \delta \leq 1$
and $t\text{-eq-}n: t = n$
and $c > 0$
and $\text{bounds}: \forall j < t. \forall \omega \in \Omega. \forall a \in A. l\text{-hat} \leq Z\text{-hat } j \ a \ \omega \wedge Z\text{-hat } j \ a \ \omega \leq$
 $u\text{-hat}$
and $\mu\text{-def}: \mu\text{-hat} = (\sum j < t. \text{expectation } (\lambda\omega. Z\text{-hat } j \ a \ \omega))$
and $u\text{-hat} - l\text{-hat} \neq 0$
and $t\text{-pos}: t > 0$
and $\forall t < n. N\text{-}n \ t \ a\text{-star } b > 0$
and $n\text{-pos}: n > 0$
and $M\text{-val } t \ a\text{-star } b \equiv (\sum s < t. \text{if } \pi \ s \ b = a\text{-star then } Z \ s \ a\text{-star } b \text{ else } 0) /$
 $\text{real } (N\text{-}n \ t \ a\text{-star } b)$
and $\text{widths}: (\sum i \in I. (b\text{-bound } i - a\text{-bound } i)^2) = (\text{real } n) * (u\text{-hat} -$
 $l\text{-hat})^2$
and $\text{space-}M: \text{space } M = \Omega$
and $\text{sets-}M: \text{sets } M = \mathcal{F}$
and $\text{indep}: \text{indep-}vars \ (\lambda\cdot. \text{borel}) \ (\lambda j. (\lambda\omega. Z \ j \ a \ \omega)) \ \{j. j < t\}$
and $\text{rv}: \forall j < t. \text{random-}variable \ \text{borel} \ (\lambda\omega. Z \ j \ a \ \omega)$
and $\forall j < t. Z\text{-hat } j \ a\text{-star } \omega = c * (\text{if } \pi \ j \ b = a\text{-star then } Z \ j \ a\text{-star } b \text{ else } 0)$
and $\text{indep-}interval\text{-bounded-}random\text{-variables } M \ I \ X\text{-new } a\text{-bound } b\text{-bound}$
and $\text{indep-}loc: \text{indep-}interval\text{-bounded-}random\text{-variables } M \ \{j. j < t\}$
 $(\lambda j. (\lambda\omega. Z\text{-hat } j \ a\text{-star } \omega))$
 $(\lambda j. l\text{-hat}) \ (\lambda j. u\text{-hat})$
and $H: \text{Hoeffding-ineq } M \ \{j. j < t\}$
 $(\lambda j. (\lambda\omega. Z\text{-hat } j \ a\text{-star } \omega))$
 $(\lambda j. l\text{-hat}) \ (\lambda j. u\text{-hat})$
and $\text{sum-}integrals\text{-eq}: (\sum j \in \{j. j < t\}. \text{integral}^L \ M \ (\lambda\omega. Z\text{-hat } j \ a\text{-star } \omega))$
 $= \mu\text{-hat}$
and $\text{rewriting}: \text{prob } \{\omega \in \Omega. (\sum j < t. Z\text{-hat } j \ a\text{-star } \omega) - (\sum j < t. \text{expectation}$
 $(Z\text{-hat } j \ a\text{-star})) \leq -\varepsilon\} =$
 $\text{prob } \{x \in \text{space } M. (\sum j < t. Z\text{-hat } j \ a\text{-star } x) \leq (\sum j < t. \text{expectation}$
 $(Z\text{-hat } j \ a\text{-star})) - \varepsilon\}$
shows $\text{prob } \{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \ a\text{-star } \omega) \leq \mu\text{-hat} - \varepsilon\}$
 $\leq \delta$
proof –

let $?I = \{j. j < t\}$
let $?X\text{-new} = \lambda j. (\lambda\omega. Z\text{-hat } j \ a\text{-star } \omega)$
let $?a\text{-bound} = \lambda j. l\text{-hat}$
let $?b\text{-bound} = \lambda j. u\text{-hat}$

have
 $M\text{-val } t \ a\text{-star } b \equiv (\sum s < t. \text{if } \pi \ s \ b = a\text{-star then } Z \ s \ a\text{-star } b \text{ else } 0) / \text{real}$
 $(N\text{-}n \ t \ a\text{-star } b)$
using assms **by** simp

have *finite-I*: *finite ?I by simp*

have *AE-bounds*: $\forall j \in ?I. AE \ \omega \text{ in } M. \ ?a\text{-bound } j \leq ?X\text{-new } j \ \omega \wedge ?X\text{-new } j \ \omega \leq ?b\text{-bound } j$

proof

fix *j* **assume** $j \in ?I$

then have $j < t$ **by** *simp*

then have *bound-j*: $\forall \omega \in \Omega. \ l\text{-hat} \leq Z\text{-hat } j \ a\text{-star } \omega \wedge Z\text{-hat } j \ a\text{-star } \omega \leq u\text{-hat}$

using *bounds by (simp add: a-in-A)*

then show *AE* ω **in** *M*. $\ ?a\text{-bound } j \leq ?X\text{-new } j \ \omega \wedge ?X\text{-new } j \ \omega \leq ?b\text{-bound } j$

using *space-M sets-M by force*

qed

have *indep-loc*: *indep-interval-bounded-random-variables M ?I ?X-new ?a-bound ?b-bound*

using *assms by simp*

have *H*: *Hoeffding-ineq M ?I ?X-new ?a-bound ?b-bound*

using *assms by simp*

have *sum-integrals-eq*: $(\sum j \in ?I. \ integral^L \ M \ (?X\text{-new } j)) = \mu\text{-hat}$

using *assms by simp*

have *widths*:

$(\sum j \in ?I. \ (?b\text{-bound } j - ?a\text{-bound } j)^2) = (\text{real } n) * (u\text{-hat} - l\text{-hat})^2$

proof –

have $(\sum j \in ?I. \ (?b\text{-bound } j - ?a\text{-bound } j)^2) =$

$(\sum j \in ?I. \ (u\text{-hat} - l\text{-hat})^2)$

by *simp*

also have $\dots = \text{card } ?I * (u\text{-hat} - l\text{-hat})^2$

by *simp*

also have $\text{card } ?I = t$

by *simp*

also have $\dots = n$

using $\langle t = n \rangle$ *assms by blast*

finally show *?thesis*

using *assms by fastforce*

qed

have *widths-pos*: $0 < (\sum j \in ?I. \ (?b\text{-bound } j - ?a\text{-bound } j)^2)$

using $\langle u\text{-hat} - l\text{-hat} \neq 0 \rangle$ *widths n-pos by auto*

have *res*: $\text{prob } \{\omega \in \Omega. \ (\sum j < t. \ Z\text{-hat } j \ a\text{-star } \omega) \leq (\sum j < t. \ \text{expectation } j)\}$

$(Z\text{-hat } j \text{ a-star}) - \varepsilon\} =$
 $\text{prob } \{x \in \text{space } M. (\sum j < t. Z\text{-hat } j \text{ a-star } x) \leq (\sum j < t. \text{expectation}$
 $(Z\text{-hat } j \text{ a-star}) - \varepsilon\}$
using *assms* **by** *simp*

then have *tail*: $\text{prob } \{\omega \in \Omega. (\sum j \in ?I. ?X\text{-new } j \omega) \leq (\sum j \in ?I. \text{integral}^L M$
 $(?X\text{-new } j)) - \varepsilon\}$
 $\leq \exp(-2 * \varepsilon^2 / (\sum j \in ?I. (?b\text{-bound } j - ?a\text{-bound } j)^2))$
using *Hoeffding-ineq.Hoeffding-ineq-le[OF H eps-pos widths-pos]*
by (*smt (verit, best) Collect-cong assms*)

have *lhs-rewrite*: $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a-star } \omega) \leq \mu\text{-hat} - \varepsilon\}$
 $= \{\omega \in \Omega. (\sum j \in ?I. ?X\text{-new } j \omega) - (\sum j \in ?I. \text{integral}^L M$
 $(?X\text{-new } j)) \leq -\varepsilon\}$
by (*metis (mono-tags, lifting) add-le-iff assms(24,6) lessThan-def*)
have *rhs-rewrite*: $\exp(-2 * \varepsilon^2 / (\sum j \in ?I. (?b\text{-bound } j - ?a\text{-bound } j)^2)) =$
 $\exp(-2 * \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2))$
using *widths* **by** *simp*

have *lhs-prob*: $\text{prob } \{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a-star } \omega) \leq \mu\text{-hat} - \varepsilon\}$
 $= \text{prob } \{\omega \in \Omega. (\sum j \in ?I. ?X\text{-new } j \omega) - (\sum j \in ?I. \text{integral}^L M$
 $(?X\text{-new } j)) \leq -\varepsilon\}$
using *lhs-rewrite* **by** *simp*

then have *prob* $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a-star } \omega) \leq \mu\text{-hat} - \varepsilon\} \leq$
 $\exp(-2 * \varepsilon^2 / (\sum j \in ?I. (?b\text{-bound } j - ?a\text{-bound } j)^2))$
using *lhs-rewrite tail lhs-prob* **by** (*smt (verit, ccfv-threshold) Collect-cong*)
then have *prob* $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a-star } \omega) \leq \mu\text{-hat} - \varepsilon\} \leq \exp(-2$
 $* \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2))$
using *rhs-rewrite* **by** *linarith*

then have *prob* $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a-star } \omega) - \mu\text{-hat} \leq -\varepsilon\} =$
 $\text{prob } \{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a-star } \omega) - \mu\text{-hat} \leq -\text{abs}(u\text{-hat} - l\text{-hat}) * \text{sqrt}$
 $((\text{real } n) / 2) * \ln(1 / \delta)\}$
using *eps* **by** *simp*

have *res1*: $\exp(-2 * \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) =$
 $\exp(-2 * (\text{abs}(u\text{-hat} - l\text{-hat}) * \text{sqrt}((\text{real } n) / 2) * \ln(1 / \delta)))^2 /$
 $(\text{real } n * (u\text{-hat} - l\text{-hat})^2)$
using *eps* **by** *simp*

have $\exp(-2 * (\text{abs}(u\text{-hat} - l\text{-hat}) * \text{sqrt}((\text{real } n) / 2) * \ln(1 / \delta)))^2 /$
 $(\text{real } n * (u\text{-hat} - l\text{-hat})^2) =$
 $\exp(-2 * (\text{abs}(u\text{-hat} - l\text{-hat}) * \text{sqrt}((\text{real } n) / 2) * \ln(1 / \delta)))^2 / (\text{real } n *$
 $(u\text{-hat} - l\text{-hat})^2)$
using *eps eps-pos* **by** *blast*

then have ... =
 $\text{exp } (- 2 * ((\text{abs } (u\text{-hat} - l\text{-hat}))^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) * \text{sqrt } (((\text{real } n / 2) * \ln (1 / \delta)))^2)$
by (*metis* (*no-types*, *opaque-lifting*) *more-arith-simps*(11) *power-mult-distrib* *times-divide-eq-left* *times-divide-eq-right*)

have ... =
 $\text{exp } (- 2 * ((\text{abs } (u\text{-hat} - l\text{-hat}))^2 / (\text{real } n * \text{abs } ((u\text{-hat} - l\text{-hat})^2)) * \text{sqrt } (((\text{real } n / 2) * \ln (1 / \delta)))^2)$
by *simp*
have ... = $\text{exp } (- 2 * (1 / (\text{real } n))) * \text{sqrt } (((\text{real } n / 2) * \ln (1 / \delta)))^2)$
using *assms* **by** *simp*

then have ... = $\text{exp } (-2 * (1 / (\text{real } n))) * (((\text{real } n / 2) * \ln (1 / \delta)))$
using *power2-eq-square* **by** (*smt* (*verit*, *best*)
arithmetic-simps(51) *assms*(10) *eps eps-pos* *real-sqrt-ge-0-iff* *real-sqrt-pow2* *zero-le-mult-iff*)

then have ... = $\text{exp } (-1 * \ln (1 / \delta))$
using *assms* **by** (*simp* *add: field-simps*)
then have fin: $\text{exp } (- 2 * \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) = \delta$
using *delta-pos* *assms*
by (*metis* (*lifting*)
 $\langle \text{exp } (- 2 * (1 / \text{real } n) * (\text{sqrt } (\text{real } n / 2 * \ln (1 / \delta)))^2) = \text{exp } (- 2 * (1 / \text{real } n) * (\text{real } n / 2 * \ln (1 / \delta))) \rangle$
 $\langle \text{exp } (- 2 * (|u\text{-hat} - l\text{-hat}| * \text{sqrt } (\text{real } n / 2 * \ln (1 / \delta)))^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) = \text{exp } (- 2 * (|u\text{-hat} - l\text{-hat}|^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) * (\text{sqrt } (\text{real } n / 2 * \ln (1 / \delta)))^2 \rangle$
 $\langle \text{exp } (- 2 * (|u\text{-hat} - l\text{-hat}|^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) * (\text{sqrt } (\text{real } n / 2 * \ln (1 / \delta)))^2) = \text{exp } (- 2 * (|u\text{-hat} - l\text{-hat}|^2 / (\text{real } n * |u\text{-hat} - l\text{-hat}|^2)) * (\text{sqrt } (\text{real } n / 2 * \ln (1 / \delta)))^2 \rangle$
 $\langle \text{exp } (- 2 * (|u\text{-hat} - l\text{-hat}|^2 / (\text{real } n * |u\text{-hat} - l\text{-hat}|^2)) * (\text{sqrt } (\text{real } n / 2 * \ln (1 / \delta)))^2) = \text{exp } (- 2 * (1 / \text{real } n) * (\text{sqrt } (\text{real } n / 2 * \ln (1 / \delta)))^2 \rangle$
arith-simps(56) *exp-ln ln-divide-pos* *ln-one* *more-arith-simps*(10) *mult-minus1* *of-nat-zero-less-power-iff* *power-0*)

show ?thesis using *assms* *fin*
 $\langle \text{prob } \{\omega \in \Omega. (\sum_{j < n} Z\text{-hat } j \text{ a-star } \omega) \leq \mu\text{-hat} - \varepsilon\} \leq \text{exp } (- 2 * \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) \rangle$
by *presburger*

qed

theorem *hoeffding-iid-bound-ge:*

fixes *a* :: 'a **and** *n* :: nat **and** ε :: real **and** $\mu\text{-hat}$:: real **and** *l-hat* *u-hat* :: real
and *I* :: nat set
and *X-new* :: nat \Rightarrow 'b \Rightarrow real

and $a\text{-bound } b\text{-bound} :: \text{nat} \Rightarrow \text{real}$
assumes $a\text{-in}: a \in A$
and $b \in \Omega$
and $\text{eps-pos}: \varepsilon \geq 0$
and $\text{eps}: \varepsilon = \text{abs } (u\text{-hat} - l\text{-hat}) * \text{sqrt } (((\text{real } n) / 2) * \ln (1 / \delta))$
and $\delta \geq 0 \wedge \delta \leq 1$
and $t\text{-eq-}n: t = n$
and $c > 0$
and $\text{bounds}: \forall j < t. \forall \omega \in \Omega. \forall a \in A. l\text{-hat} \leq Z\text{-hat } j \ a \ \omega \wedge Z\text{-hat } j \ a \ \omega \leq$
 $u\text{-hat}$
and $\mu\text{-def}: \mu\text{-hat} = (\sum j < t. \text{expectation } (\lambda \omega. Z\text{-hat } j \ a \ \omega))$
and $u\text{-hat} - l\text{-hat} \neq 0$
and $t\text{-pos}: t > 0$
and $\forall t < n. N\text{-}n \ t \ a\text{-star } b > 0$
and $n\text{-pos}: n > 0$
and $M\text{-val } t \ a \ b \equiv (\sum s < t. \text{if } \pi \ s \ b = a \ \text{then } Z \ s \ a \ b \ \text{else } 0) /$
 $\text{real } (N\text{-}n \ t \ a \ b)$
and $\text{widths}: (\sum i \in I. (b\text{-bound } i - a\text{-bound } i)^2) = (\text{real } n) * (u\text{-hat} -$
 $l\text{-hat})^2$
and $\text{space-}M: \text{space } M = \Omega$
and $\text{sets-}M: \text{sets } M = \mathcal{F}$
and $\text{indep}: \text{indep-}vars \ (\lambda\cdot. \text{borel}) \ (\lambda j. (\lambda \omega. Z \ j \ a \ \omega)) \ \{j. j < t\}$
and $\text{rv}: \forall j < t. \text{random-}variable \ \text{borel} \ (\lambda \omega. Z \ j \ a \ \omega)$
and $\forall j < t. Z\text{-hat } j \ a \ \omega = c * (\text{if } \pi \ j \ b = a \ \text{then } Z \ j \ a\text{-star } b \ \text{else } 0)$
and $\text{indep-}interval\text{-}bounded\text{-}random\text{-}variables \ M \ I \ X\text{-new } a\text{-bound } b\text{-bound}$
and $\text{indep-}loc: \text{indep-}interval\text{-}bounded\text{-}random\text{-}variables \ M \ \{j. j < t\}$
 $(\lambda j. (\lambda \omega. Z\text{-hat } j \ a \ \omega))$
 $(\lambda j. l\text{-hat}) \ (\lambda j. u\text{-hat})$
and $H: \text{Hoeffding-}ineq \ M \ \{j. j < t\}$
 $(\lambda j. (\lambda \omega. Z\text{-hat } j \ a \ \omega))$
 $(\lambda j. l\text{-hat}) \ (\lambda j. u\text{-hat})$
and $\text{sum-}integrals\text{-}eq: (\sum j \in \{j. j < t\}. \text{integral}^L \ M \ (\lambda \omega. Z\text{-hat } j \ a \ \omega)) =$
 $\mu\text{-hat}$
and $\text{rewriting}: \text{prob } \{\omega \in \Omega. (\sum j < t. Z\text{-hat } j \ a \ \omega) - (\sum j < t. \text{expectation}$
 $(Z\text{-hat } j \ a)) \geq \varepsilon\} =$
 $\text{prob } \{x \in \text{space } M. (\sum j < t. Z\text{-hat } j \ a \ x) \geq (\sum j < t. \text{expectation } (Z\text{-hat } j$
 $a)) + \varepsilon\}$
shows $\text{prob } \{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \ a \ \omega) \geq \mu\text{-hat} + \varepsilon\} \leq \delta$
proof –

let $?I = \{j. j < t\}$
let $?X\text{-new} = \lambda j. (\lambda \omega. Z\text{-hat } j \ a \ \omega)$
let $?a\text{-bound} = \lambda j. l\text{-hat}$
let $?b\text{-bound} = \lambda j. u\text{-hat}$

have
 $M\text{-val } t \ a \ b \equiv (\sum s < t. \text{if } \pi \ s \ b = a \ \text{then } Z \ s \ a \ b \ \text{else } 0) / \text{real } (N\text{-}n \ t \ a \ b)$
using assms **by** simp

have *finite-I*: *finite ?I by simp*

have *AE-bounds*: $\forall j \in ?I. AE \ \omega \text{ in } M. \ ?a\text{-bound } j \leq ?X\text{-new } j \ \omega \wedge ?X\text{-new } j \ \omega \leq ?b\text{-bound } j$

proof

fix *j* **assume** $j \in ?I$

then have $j < t$ **by** *simp*

then have *bound-j*: $\forall \omega \in \Omega. l\text{-hat} \leq Z\text{-hat } j \ a \ \omega \wedge Z\text{-hat } j \ a \ \omega \leq u\text{-hat}$

using *bounds* **by** (*simp add: a-in-A*)

then show *AE* $\omega \text{ in } M. \ ?a\text{-bound } j \leq ?X\text{-new } j \ \omega \wedge ?X\text{-new } j \ \omega \leq ?b\text{-bound } j$

using *space-M sets-M* **by** *force*

qed

have *indep-loc*: *indep-interval-bounded-random-variables* $M \ ?I \ ?X\text{-new} \ ?a\text{-bound} \ ?b\text{-bound}$

using *assms* **by** *simp*

have *H*: *Hoeffding-ineq* $M \ ?I \ ?X\text{-new} \ ?a\text{-bound} \ ?b\text{-bound}$

using *assms* **by** *simp*

have *sum-integrals-eq*: $(\sum_{j \in ?I. \ integral^L \ M \ (?X\text{-new } j)) = \mu\text{-hat}$

using *assms* **by** *simp*

have *widths*:

$(\sum_{j \in ?I. \ (?b\text{-bound } j - ?a\text{-bound } j)^2) = (real \ n) * (u\text{-hat} - l\text{-hat})^2$

proof –

have $(\sum_{j \in ?I. \ (?b\text{-bound } j - ?a\text{-bound } j)^2) =$

$(\sum_{j \in ?I. \ (u\text{-hat} - l\text{-hat})^2)$

by *simp*

also have $\dots = card \ ?I * (u\text{-hat} - l\text{-hat})^2$

by *simp*

also have $card \ ?I = t$

by *simp*

also have $\dots = n$

using $\langle t = n \rangle$ *assms* **by** *blast*

finally show *?thesis*

using *assms* **by** *fastforce*

qed

have *widths-pos*: $0 < (\sum_{j \in ?I. \ (?b\text{-bound } j - ?a\text{-bound } j)^2)$

using $\langle u\text{-hat} - l\text{-hat} \neq 0 \rangle$ *widths n-pos* **by** *auto*

have *res*: $prob \ \{\omega \in \Omega. \ (\sum_{j < t. \ Z\text{-hat } j \ a \ \omega) - (\sum_{j < t. \ expectation \ (Z\text{-hat } j \ a)) \geq \varepsilon\} =$

$\text{prob } \{x \in \text{space } M. (\sum j < t. Z\text{-hat } j \text{ a } x) \geq (\sum j < t. \text{expectation } (Z\text{-hat } j \text{ a})) + \varepsilon\}$

using *assms* **by** *simp*

then have *tail*: $\text{prob } \{\omega \in \Omega. (\sum j \in ?I. ?X\text{-new } j \omega) \geq (\sum j \in ?I. \text{integral}^L M (\text{?X-new } j)) + \varepsilon\}$

$\leq \exp(-2 * \varepsilon^2 / (\sum j \in ?I. (?b\text{-bound } j - ?a\text{-bound } j)^2))$

using *Hoeffding-ineq.Hoeffding-ineq-ge[OF H eps-pos widths-pos]*

by (*smt (verit, best) Collect-cong assms*)

have *lhs-rewrite*: $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a } \omega) \geq \mu\text{-hat} + \varepsilon\}$

$= \{\omega \in \Omega. (\sum j \in ?I. ?X\text{-new } j \omega) - (\sum j \in ?I. \text{integral}^L M$

$(\text{?X-new } j)) \geq \varepsilon\}$

using *assms* **by** (*metis (lifting) ext add commute le-diff-eq lessThan-def*)

have *rhs-rewrite*: $\exp(-2 * \varepsilon^2 / (\sum j \in ?I. (?b\text{-bound } j - ?a\text{-bound } j)^2)) = \exp(-2 * \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2))$

using *widths* **by** *simp*

have *lhs-prob*: $\text{prob } \{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a } \omega) \geq \mu\text{-hat} + \varepsilon\}$

$= \text{prob}\{\omega \in \Omega. (\sum j \in ?I. ?X\text{-new } j \omega) - (\sum j \in ?I. \text{integral}^L M$

$(\text{?X-new } j)) \geq \varepsilon\}$

using *assms lhs-rewrite* **by** *presburger*

then have *prob* $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a } \omega) \geq \mu\text{-hat} + \varepsilon\} \leq$

$\exp(-2 * \varepsilon^2 / (\sum j \in ?I. (?b\text{-bound } j - ?a\text{-bound } j)^2))$

using *lhs-rewrite tail lhs-prob* **by** (*smt (verit, ccfv-threshold) Collect-cong*)

then have *prob* $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a } \omega) \geq \mu\text{-hat} + \varepsilon\} \leq \exp(-2 * \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2))$

using *rhs-rewrite* **by** *linarith*

then have *prob* $\{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a } \omega) - \mu\text{-hat} \geq \varepsilon\} =$

$\text{prob } \{\omega \in \Omega. (\sum j < n. Z\text{-hat } j \text{ a } \omega) - \mu\text{-hat} \geq \text{abs } (u\text{-hat} - l\text{-hat}) * \text{sqrt } (((\text{real } n) / 2) * \ln(1 / \delta))\}$

using *eps* **by** *simp*

have *res1*: $\exp(-2 * \varepsilon^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) =$

$\exp(-2 * (\text{abs } (u\text{-hat} - l\text{-hat}) * \text{sqrt } ((\text{real } n / 2) * \ln(1 / \delta)))^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2))$

using *eps* **by** *simp*

have $\exp(-2 * (\text{abs } (u\text{-hat} - l\text{-hat}) * \text{sqrt } ((\text{real } n / 2) * \ln(1 / \delta)))^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2)) =$

$\exp(-2 * (\text{abs } (u\text{-hat} - l\text{-hat}) * \text{sqrt } ((\text{real } n / 2) * \ln(1 / \delta)))^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2))$

using *eps eps-pos* **by** *blast*

then have ... $= \exp(-2 * ((\text{abs } (u\text{-hat} - l\text{-hat}))^2 / (\text{real } n * (u\text{-hat} - l\text{-hat})^2))$

```

)) * sqrt (((real n / 2) * ln (1 / δ)))^2 )
  by (metis (no-types, opaque-lifting) more-arith-simps(11) power-mult-distrib
times-divide-eq-left
      times-divide-eq-right)

  also have ... = exp (- 2 * ((abs (u-hat - l-hat))^2 / (real n * abs ((u-hat -
l-hat))^2 )) * sqrt (((real n / 2) * ln (1 / δ)))^2 )
  by simp
  have ... = exp (- 2 * (1 / (real n) )) * sqrt (((real n / 2) * ln (1 / δ)))^2 )
  using assms by simp

  then have ... = exp (-2 * (1 / (real n) )) * (((real n / 2) * ln (1 / δ))) )
  using power2-eq-square by (smt (verit, best)
arithmetic-simps(51) assms(10) eps eps-pos real-sqrt-ge-0-iff real-sqrt-pow2
zero-le-mult-iff)

  then have ... = exp (-1 * ln (1 / δ))
  using assms by (simp add: field-simps)
  then have fin: exp (- 2 * ε^2 / (real n * (u-hat - l-hat)^2)) = δ
  using δ-pos assms
  by (metis (lifting)
      ⟨exp (- 2 * (1 / real n) * (sqrt (real n / 2 * ln (1 / δ)))^2) = exp (- 2 *
(1 / real n) * (real n / 2 * ln (1 / δ)))⟩
      ⟨exp (- 2 * (|u-hat - l-hat| * sqrt (real n / 2 * ln (1 / δ)))^2 / (real n *
(u-hat - l-hat)^2)) = exp (- 2 * (|u-hat - l-hat|^2 / (real n * (u-hat - l-hat)^2)) *
(sqrt (real n / 2 * ln (1 / δ)))^2)⟩
      ⟨exp (- 2 * (|u-hat - l-hat|^2 / (real n * (u-hat - l-hat)^2)) * (sqrt (real n /
2 * ln (1 / δ)))^2) = exp (- 2 * (|u-hat - l-hat|^2 / (real n * |u-hat - l-hat|^2)) *
(sqrt (real n / 2 * ln (1 / δ)))^2)⟩
      ⟨exp (- 2 * (|u-hat - l-hat|^2 / (real n * |u-hat - l-hat|^2)) * (sqrt (real n /
2 * ln (1 / δ)))^2) = exp (- 2 * (1 / real n) * (sqrt (real n / 2 * ln (1 / δ)))^2)⟩
      diff-0 exp-ln ln-divide-pos ln-one more-arith-simps(10) mult-minus1 of-nat-zero-less-power-iff
power-0)

  show ?thesis using assms fin
  using ⟨prob {ω ∈ Ω. μ-hat + ε ≤ (∑ j<n. Z-hat j a ω)} ≤ exp (- 2 * ε^2 /
(real n * (u-hat - l-hat)^2))⟩
  by presburger

qed

end
end
theory Discrete-UCB-Step3
  imports Discrete-UCB-Step2

begin

```

locale *bandit-policy* = *bandit-problem* + *prob-space* +
fixes
 $\Omega :: 'b \text{ set}$
and $\mathcal{F} :: 'b \text{ set set}$
and $\pi :: \text{nat} \Rightarrow 'b \Rightarrow 'a$
and $N\text{-}n :: \text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{nat}$
and $Z :: \text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$
and $\delta :: \text{real}$
and $q :: \text{real}$
assumes *fin-A*: *finite A*
and $\omega \in \Omega$
and *a-in-A*: $a \in A$
and *measurable-policy*: $\forall t. \pi \ t \in \text{measurable } M \ (\text{count-space } A)$
and *N-n-def*: $\forall n \ a \ \omega. N\text{-}n \ n \ a \ \omega = \text{card } \{t \in \{0..<n\}. \pi \ (t+1) \ \omega = a\}$
and *count-assm-pointwise*: $\forall n \ \omega. (\sum a \in A. \text{real } (N\text{-}n \ n \ a \ \omega)) = \text{real } n$
and *delta-pos*: $0 < \delta$
and *delta-less1*: $\delta < 1$
and *q-pos*: $q > 0$

begin

definition *sample-mean-Z* :: $\text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$ **where**
sample-mean-Z $n \ a \ \omega = (1 / \text{real } n) * (\sum i < n. Z \ i \ a \ \omega)$

definition *M-fun* :: $\text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$ **where**
M-fun $t \ a \ \omega = (\text{if } N\text{-}n \ (t+1) \ a \ \omega = 0 \ \text{then } 0$
else $(\sum s < t. (\text{if } \pi \ s \ \omega = a \ \text{then } Z \ s \ a \ \omega \ \text{else } 0)) / \text{real } (N\text{-}n \ t \ a \ \omega))$

definition *U* :: $\text{nat} \Rightarrow 'a \Rightarrow 'b \Rightarrow \text{real}$ **where**
U $t \ a \ \omega = M\text{-}fun \ t \ a \ \omega + q * \text{sqrt } (\ln \ (1 / \delta) / (2 * \text{real } (\max \ 1 \ (N\text{-}n \ t \ a \ \omega))))$

definition *A-t-plus-1* :: $\text{nat} \Rightarrow 'b \Rightarrow 'a$ **where**
A-t-plus-1 $t \ \omega = (\text{SOME } a. a \in A \wedge (\forall a'. a' \in A \longrightarrow U \ t \ a \ \omega \geq U \ t \ a' \ \omega))$

definition *prob-eq-Ex* :: $'b \text{ set} \Rightarrow \text{bool}$ **where**
prob-eq-Ex $E \equiv \text{prob } E = \text{expectation } (\lambda \omega. \text{indicator } E \ \omega)$

theorem *proposition-15-7*:
assumes
a-in-A: $a \in A$
 $\omega \in \Omega$
and *subopt-gap*: $\Delta \ a > 0$
and *a-not-opt*: $\exists a' \in A. \Delta \ a' > 0$
and *delta-a*: $\forall a \in A. \Delta \ a = \mu \ a\text{-star} - \mu \ a$
and $k \leq n$
and *from-UCB-step1*: $\forall a \in A. \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega)) \leq$

$s n + \text{expectation } (\lambda\omega. (\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge s t \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0))$

and from-UCB-step2: $\forall a \in A. \forall t \in \{k..<n\}. \text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap$

$$\{\omega \in \Omega. N\text{-}n \text{ } t a \omega \geq (2 * \ln (1 / \delta)) / (\Delta a)^{\wedge 2}\}) \leq 2 * \delta$$

and eps-pos: $\varepsilon > 0$

and t-gt0: $\forall t \in \{k..<n\}. t > 0$

and choice-delta: $\forall t \in \{k..<n\}. \delta = 1 / (\text{real } t \text{ powr } \varepsilon)$

and s-form: $\forall a \in A. \forall u. s u = (2 * \varepsilon * \ln (\text{real } u)) / ((\Delta a)^{\wedge 2})$

and subset-meas: $\forall a \in A. \forall t \in \{k..<n\}. \forall \omega \in \Omega. \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq N\text{-}n \text{ } t a \omega\} \subseteq \Omega$

and prob-eq-E-assm: $\forall a \in A. \forall t \in \{k..<n\}. \text{prob } \{\omega. \pi (Suc t) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \text{ } t a \omega)\} =$

$\text{prob } (\{\omega. \pi (Suc t) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \text{ } t a \omega)\} \cap \text{space } M)$

and finiteness: $\forall t \in \{k..<n\}. \forall a \in A. \text{emeasure } M \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \text{ } t a \omega)\} < \infty$

and measurable-set: $\forall t \in \{k..<n\}. \forall a \in A. \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \text{ } t a \omega)\} \in \text{sets } M$

and eq-sets-optimum:

$\forall a \in A. \forall t \in \{k..<n\}. \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \text{ } t a \omega)\} =$

$\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. N\text{-}n \text{ } t a \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}\}$

shows

$\forall a \in A. \text{expectation } (\lambda\omega. \text{real } (N\text{-}n \text{ } n a \omega)) \leq s n + (\sum t = k..<n. 2 / (\text{real } t \text{ powr } \varepsilon))$

proof –

have def-sn: $\forall a \in A. s n = (2 * \varepsilon * \ln (\text{real } n)) / ((\Delta a)^{\wedge 2})$

using s-form by simp

have def-st: $\forall a \in A. s t = (2 * \varepsilon * \ln (\text{real } t)) / ((\Delta a)^{\wedge 2})$

using s-form by simp

then have expression: $\forall a \in A. \text{expectation } (\lambda\omega. \text{real } (N\text{-}n \text{ } n a \omega)) \leq$

$(2 * \varepsilon * \ln (\text{real } n)) / ((\Delta a)^{\wedge 2}) + \text{expectation } (\lambda\omega. (\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge$

$(2 * \varepsilon * \ln (\text{real } t)) / ((\Delta a)^{\wedge 2}) \leq \text{real } (N\text{-}n \text{ } t a \omega) \text{ then } 1 \text{ else } 0))$

using assms def-sn def-st by simp

have eq-if-of-bool:

$\forall a \in A. \text{expectation } (\lambda\omega. \sum t = k..<n.$

$(\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \text{ } t a \omega) \text{ then } 1 \text{ else } 0))$

$= \text{expectation } (\lambda\omega. \sum t = k..<n.$

of-bool ($\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)$)
by (simp add: of-bool-def)

have eq-indic-bool:

$\forall a \in A. \text{expectation } (\lambda \omega. \sum t = k..<n. \text{of-bool } (\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)))$
 $= \text{expectation } (\lambda \omega. \sum t = k..<n. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} \omega)$
by (simp add: indicator-def of-bool-def)

have expression-1:

$\forall a \in A. \text{expectation } (\lambda \omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$
 $\text{expectation } (\lambda \omega. \sum t = k..<n. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} \omega)$

proof

fix a **assume** $a \in A$

from eq-if-of-bool[rule-format, OF $\langle a \in A \rangle$]

have eq1: $\text{expectation } (\lambda \omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$

$\text{expectation } (\lambda \omega. \sum t = k..<n. \text{of-bool } (\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)))$
by simp

from eq-indic-bool[rule-format, OF $\langle a \in A \rangle$]

have eq2: $\text{expectation } (\lambda \omega. \sum t = k..<n. \text{of-bool } (\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega))) =$

$\text{expectation } (\lambda \omega. \sum t = k..<n. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} \omega)$
by simp

from eq1 eq2 **show**

$\text{expectation } (\lambda \omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$

$\text{expectation } (\lambda \omega. \sum t = k..<n. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} \omega)$
by (rule trans)

qed

have res: $\forall a \in A. \forall t \in \{k..<n\}. \text{prob-eq-Ex } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$

using *assms prob-eq-Ex-def* **by** *auto*

have *lin-of-expect-indicators*:

$\forall a \in A. \forall t \in \{k..<n\}. \text{prob } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\} =$
 $\text{expectation } (\lambda\omega. (\text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\} \omega))$

using *prob-eq-Ex-def res* **by** *simp*

then have *key-result-1*: $\forall a \in A. (\sum t = k..<n.$

$\text{expectation } (\lambda\omega. (\text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\} \omega))$
 $\leq \text{real } (N-n t a \omega) \omega))$
 $= (\sum t = k..<n. \text{prob } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\})$

using *lin-of-expect-indicators* **by** *simp*

have *expression-follow-up*: $\forall a \in A.$

$\text{expectation } (\lambda\omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$
 $(\sum t = k..<n. \text{prob } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\})$

proof –

have *res1*: $\forall a \in A.$

$\text{expectation } (\lambda\omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$
 $\text{expectation } (\lambda\omega. \sum t = k..<n.$

$\text{of-bool } (\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)))$

using *eq-if-of-bool* **by** *simp*

have *res2*: $\forall a \in A.$

$\text{expectation } (\lambda\omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$
 $\text{expectation } (\lambda\omega. \sum t = k..<n. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\} \omega)$

using *expression-1* **by** *simp*

have *res3*: $\forall a \in A.$

$\text{expectation } (\lambda\omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$
 $(\sum t = k..<n. \text{expectation } (\lambda\omega. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\} \omega))$

proof –

have $\forall a \in A.$

$\text{expectation } (\lambda\omega. \sum t = k..<n. (\text{if } \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0)) =$
 $\text{expectation } (\lambda\omega. \sum t = k..<n. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^2 \leq \text{real } (N-n t a \omega)\} \omega)$

using *res2* **by** *blast*

then have $\forall a \in A.$ *expectation* $(\lambda\omega. \sum t = k..<n.$ *indicat-real* $\{\omega. \pi (t+1)$
 $\omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega) =$
integral^L $M (\lambda\omega. \sum t = k..<n.$ *indicat-real* $\{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega)$

by simp

then have *result-intermed*: $\forall a \in A.$ *expectation* $(\lambda\omega. \sum t = k..<n.$ *indicat-real* $\{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega) =$
 $(\sum t = k..<n.$ *integral*^L $M (\lambda\omega. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega))$

using *assms integral-sum by simp*

have *result*: $\forall a \in A.$ *expectation* $(\lambda\omega. \sum t = k..<n.$ *indicat-real* $\{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega) =$
 $(\sum t = k..<n.$ *expectation* $(\lambda\omega. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega))$

using *result-intermed by auto*

have $\forall a \in A.$

expectation $(\lambda\omega. \sum t = k..<n.$ *(if* $\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)$ *then 1 else 0)) =*
expectation $(\lambda\omega. \sum t = k..<n.$ *indicat-real* $\{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega)$

using *res2 by simp*

then have *linearity-expectation*: $\forall a \in A.$

expectation $(\lambda\omega. \sum t = k..<n.$ *(if* $\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)$ *then 1 else 0)) =*
 $(\sum t = k..<n.$ *expectation* $(\lambda\omega. \text{indicat-real } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}$ $\omega))$

using *result by simp*

then show *?thesis using linearity-expectation by simp*

qed

have *final-linearity*: $\forall a \in A.$

expectation $(\lambda\omega. \sum t = k..<n.$ *(if* $\pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)$ *then 1 else 0)) =*
 $(\sum t = k..<n.$ *prob* $\{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\})$

using *res1 res2 res3 key-result-1 by simp*

then show *?thesis using final-linearity by simp*

qed

have *intermed-result*: $\forall a \in A.$

expectation $(\lambda\omega. \text{real } (N-n n a \omega)) \leq$
 $(2 * \varepsilon * \ln (\text{real } n)) / (\Delta a)^{\wedge 2} +$
 $(\sum t = k..<n.$ *prob* $\{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\})$

using *expression-follow-up expression by simp*

then have follow-up-result: $\forall a \in A. \forall t \in \{k..<n\}.$
 $\text{prob } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} =$
 $\text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. N\text{-}n t a \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}\})$
using expression-follow-up expression eq-sets-optimum intermed-result by simp

have next-result-sum-prob: $\forall a \in A.$
 $(\sum t = k..<n. \text{prob } \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\}) =$
 $(\sum t = k..<n. \text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. N\text{-}n t a \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}\}))$
using assms follow-up-result by simp

then have next-result-fin: $\forall a \in A.$
 $\text{expectation } (\lambda \omega. \text{real } (N-n n a \omega)) \leq$
 $(2 * \varepsilon * \ln (\text{real } n)) / (\Delta a)^{\wedge 2} +$
 $(\sum t = k..<n. \text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. N\text{-}n t a \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}\}))$
using expression-follow-up next-result-sum-prob expression by fastforce

have generalized-bound:
 $\forall a \in A. \forall t \in \{k..<n\}. \text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap$
 $\{\omega \in \Omega. N\text{-}n t a \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}\}) \leq 2 /$
 $(\text{real } t \text{ powr } \varepsilon)$

proof
fix a
assume $a\text{-in}: a \in A$
show $\forall t \in \{k..<n\}. \text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. N\text{-}n t a \omega$
 $\geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}\}) \leq 2 / (\text{real } t \text{ powr } \varepsilon)$

proof
fix t
assume $t\text{-in}: t \in \{k..<n\}$
have $H\text{prob}:$
 $\text{prob } (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap$
 $\{\omega \in \Omega. N\text{-}n t a \omega \geq (2 * \ln (1 / \delta)) / (\Delta a)^{\wedge 2}\}) \leq 2 * \delta$
using from-UCB-step2[rule-format, OF a-in t-in] by blast

have $\ln\text{-eq}: \ln (1 / \delta) = \varepsilon * \ln (\text{real } t)$
using choice-delta[rule-format, OF t-in] eps-pos by simp

hence $\text{threshold-eq}:$
 $(2 * \ln (1 / \delta)) / (\Delta a)^{\wedge 2} = (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}$
by simp

hence $\text{set-eq}:$
 $\{\omega \in \Omega. N\text{-}n t a \omega \geq (2 * \ln (1 / \delta)) / (\Delta a)^{\wedge 2}\} =$
 $\{\omega \in \Omega. N\text{-}n t a \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta a)^{\wedge 2}\}$

by auto

from *Hprob set-eq* **have**
 $\text{prob} (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \ \omega = a\} \cap$
 $\{\omega \in \Omega. N\text{-}n \ t \ a \ \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta \ a)^2\}) \leq 2 * \delta$
by simp

moreover from *choice-delta*[*rule-format*, *OF t-in*] **have** $2 * \delta = 2 / (\text{real } t$
powr $\varepsilon)$
by simp

ultimately show
 $\text{prob} (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \ \omega = a\} \cap$
 $\{\omega \in \Omega. N\text{-}n \ t \ a \ \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta \ a)^2\}) \leq 2 / (\text{real } t \ \text{powr}$
 $\varepsilon)$
by simp

qed
qed

have *sum-mono-expression*:
 $\forall a \in A. (\sum t = k..<n. \text{prob} (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \ \omega = a\} \cap$
 $\{\omega \in \Omega. N\text{-}n \ t \ a \ \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta \ a)^2\}))$
 $\leq (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))$
proof
fix a
assume *a-in*: $a \in A$
 $\text{show } (\sum t = k..<n. \text{prob} (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \ \omega = a\} \cap \{\omega \in \Omega. N\text{-}n \ t \ a \ \omega$
 $\geq (2 * \varepsilon * \ln (\text{real } t)) / (\Delta \ a)^2\}))$
 $\leq (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))$
using *generalized-bound a-in* **by** (*intro sum-mono*) *auto*
qed

have *final-result*:
 $\forall a \in A. \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega))$
 $\leq (2 * \varepsilon * \ln (\text{real } n)) / ((\Delta \ a)^2) + (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))$
proof
fix a
assume *a-in*: $a \in A$
from *next-result-fin*[*rule-format*, *OF a-in*]
have *bound1*:
 $\text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega))$
 $\leq (2 * \varepsilon * \ln (\text{real } n)) / ((\Delta \ a)^2)$
 $+ (\sum t = k..<n. \text{prob} (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \ \omega = a\}$
 $\cap \{\omega \in \Omega. N\text{-}n \ t \ a \ \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / ((\Delta$
 $a)^2\}))$
by simp
moreover from *sum-mono-expression*[*rule-format*, *OF a-in*]
have *bound2*:
 $(\sum t = k..<n. \text{prob} (\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \ \omega = a\}$

$\cap \{\omega \in \Omega. N-n t a \omega \geq (2 * \varepsilon * \ln (\text{real } t)) / ((\Delta a)^{\wedge 2})\}$
 $\leq (\sum t = k..<n. 2 / (\text{real } t \text{ powr } \varepsilon))$
by simp
ultimately show
 $\text{expectation } (\lambda \omega. \text{real } (N-n n a \omega))$
 $\leq (2 * \varepsilon * \ln (\text{real } n)) / ((\Delta a)^{\wedge 2}) + (\sum t = k..<n. 2 / (\text{real } t \text{ powr } \varepsilon))$
by auto
qed

show ?thesis using assms final-result by simp

qed

theorem theorem-15-4:

assumes

a-in-A: $a \in A$

and finite A **and** $\forall a \in A. \text{integrable } M (\lambda \omega. \text{real } (N-n n a \omega))$

and $\omega\text{-in-}\Omega$: $\omega \in \Omega$

and subopt-gap: $\forall a \in A. \Delta a > 0$

and a-not-opt: $\exists a' \in A. \Delta a' > 0$

and delta-a: $\forall a \in A. \Delta a = \mu a\text{-star} - \mu a$

and $k \leq n$

and n-count-assm-pointwise: $(\sum a \in A. \text{real } (N-n n a \omega)) = \text{real } n$

and expected-regret-prop-15-1: $\text{expectation } (\lambda \omega. R-n n \omega) = (\sum a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N-n n a \omega)))$

and from-UCB-step1: $\forall a \in A. \text{expectation } (\lambda \omega. \text{real } (N-n n a \omega)) \leq$

$s n + \text{expectation } (\lambda \omega. (\sum t = k..<n. \text{if } \pi (t+1) \omega = a \wedge s t \leq \text{real } (N-n t a \omega) \text{ then } 1 \text{ else } 0))$

and from-UCB-step2: $\forall a \in A. \forall t \in \{k..<n\}. \text{prob } (\{\omega \in \Omega. A\text{-t-plus-1 } t \omega = a\} \cap$

$\{\omega \in \Omega. N-n t a \omega \geq (2 * \ln (1 / \delta)) / (\Delta a)^{\wedge 2}\}) \leq 2 * \delta$

and eps-pos: $\varepsilon > 0$

and t-gt0: $\forall t \in \{k..<n\}. t > 0$

and choice-delta: $\forall t \in \{k..<n\}. \delta = 1 / (\text{real } t \text{ powr } \varepsilon)$

and s-form: $\forall a \in A. \forall u. s u = (2 * \varepsilon * \ln (\text{real } u)) / ((\Delta a)^{\wedge 2})$

and subset-meas: $\forall a \in A. \forall t \in \{k..<n\}. \forall \omega \in \Omega. \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq N-n t a \omega\} \subseteq \Omega$

and prob-eq-E-assm: $\forall a \in A. \forall t \in \{k..<n\}. \text{prob } \{\omega. \pi (Suc t) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} =$

$\text{prob } (\{\omega. \pi (Suc t) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} \cap \text{space } M)$

and finiteness: $\forall t \in \{k..<n\}. \forall a \in A. \text{emeasure } M \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} < \infty$

and measurable-set: $\forall t \in \{k..<n\}. \forall a \in A. \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} \in \text{sets } M$

and eq-sets-optimum:

$\forall a \in A. \forall t \in \{k..<n\}. \{\omega. \pi (t+1) \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N-n t a \omega)\} =$

$\{\omega \in \Omega. A\text{-}t\text{-plus-1 } t \omega = a\} \cap \{\omega \in \Omega. N\text{-}n \ t \ a \ \omega \geq (2 * \varepsilon * \ln$
 $(\text{real } t)) / (\Delta a)^{\wedge 2}\}$
and *assms-lin-expect*: $\forall a \in A. \text{expectation } (\lambda \omega. \sum t = k..<n.$
 $(\text{if } \pi (t+1) \ \omega = a \wedge 2 * \varepsilon * \ln (\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \ t \ a \ \omega) \text{ then}$
 $1 \text{ else } 0)) =$
 $(\sum t = k..<n. \text{expectation } (\lambda \omega. \text{indicat-real } \{\omega. \pi (t+1) \ \omega = a \wedge 2 * \varepsilon * \ln$
 $(\text{real } t) / (\Delta a)^{\wedge 2} \leq \text{real } (N\text{-}n \ t \ a \ \omega)\} \ \omega))$
and *mono-sum-sets*:
 $(\forall a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega)) \leq \Delta a * (s \ n + (\sum t = k..<n.$
 $2 / (\text{real } t \ \text{powr } \varepsilon))))$
 $\implies (\sum a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega))) \leq$
 $(\sum a \in A. \Delta a * (s \ n + (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))))$
shows $\text{expectation } (\lambda \omega. R\text{-}n \ n \ \omega) \leq (\sum a \in A. \Delta a * ((2 * \varepsilon * \ln (\text{real } n)) /$
 $((\Delta a)^{\wedge 2}) + (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))))$
proof –
have *exp-regret-eq*:
 $\text{expectation } (\lambda \omega. R\text{-}n \ n \ \omega) = (\sum a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a$
 $\ \omega)))$
using *assms by fastforce*

have *result-2*: $\forall a \in A. \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega))$
 $\leq s \ n + (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))$
using *assms proposition-15-7 by (smt (verit, ccfv-threshold) a-star-in-A)*

have *intermed-step*: $\forall a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega)) \leq \Delta a * (s$
 $\ n + (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon)))$
proof
fix *a* **assume** $a \in A$
then **have** $\text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega)) \leq s \ n + (\sum t = k..<n. 2 /$
 $(\text{real } t \ \text{powr } \varepsilon))$
using *result-2 by auto*
moreover **have** $0 < \Delta a$
using *assms subopt-gap <a ∈ A> by blast*
ultimately show $\Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega)) \leq \Delta a * (s \ n +$
 $(\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon)))$
using *mult-left-mono by simp*
qed

have *intermed-step-2*: $(\sum a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega))) \leq$
 $(\sum a \in A. \Delta a * (s \ n + (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))))$
proof –
have *res*: $\forall a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega)) \leq \Delta a * (s \ n + (\sum$
 $t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon)))$
using *intermed-step by auto*
have *res2*: $(\forall a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega)) \leq \Delta a * (s \ n +$
 $(\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))))$
 $\implies (\sum a \in A. \Delta a * \text{expectation } (\lambda \omega. \text{real } (N\text{-}n \ n \ a \ \omega))) \leq (\sum a \in A. \Delta a * (s$
 $\ n + (\sum t = k..<n. 2 / (\text{real } t \ \text{powr } \varepsilon))))$

using *res assms* **by** *fastforce*
then have *intermed-step*: $(\sum_{a \in A} \Delta a * \text{expectation}(\lambda \omega. \text{real}(N - n \ n \ a \ \omega)))$
 $\leq (\sum_{a \in A} \Delta a * (s \ n + (\sum_{t = k..<n} 2 / (\text{real } t \ \text{powr } \varepsilon))))$
using *res res2* **by** *auto*

then show *?thesis* **using** *intermed-step* **by** *simp*
qed

then have *expectation* $(\lambda \omega. R - n \ n \ \omega) \leq (\sum_{a \in A} \Delta a * (s \ n + (\sum_{t = k..<n} 2 / (\text{real } t \ \text{powr } \varepsilon))))$
using *assms intermed-step-2* **by** *linarith*

have $\forall a \in A. s \ n = (2 * \varepsilon * \ln(\text{real } n)) / ((\Delta a)^2)$
using *s-form* **by** *auto*

then have *expectation* $(\lambda \omega. R - n \ n \ \omega) \leq (\sum_{a \in A} \Delta a * ((2 * \varepsilon * \ln(\text{real } n)) / ((\Delta a)^2) + (\sum_{t = k..<n} 2 / (\text{real } t \ \text{powr } \varepsilon))))$
using *assms intermed-step-2* **by** *(metis a-star-in-A less-irrefl verit-minus-simplify(1))*

then show *?thesis* **by** *simp*

qed

end
end

References

- [1] R. Rebeschini. Lecture 15: Stochastic multi-armed bandit problem and algorithms. 2022. Available at <https://www.stats.ox.ac.uk/~rebeschini/teaching/AFoL/22/material/lecture15.pdf>.