A Decidable Confluence Test for Cognitive Models in ACT-R

Daniel Gall and Thom Frühwirth

Institute of Software Engineering and Programming Languages, Ulm University, 89069 Ulm, Germany, {daniel.gall,thom.fruehwirth}@uni-ulm.de

Abstract. Computational cognitive modeling investigates human cognition by building detailed computational models for cognitive processes. Adaptive Control of Thought – Rational (ACT-R) is a rule-based cognitive architecture that offers a widely employed framework to build such models. There is a sound and complete embedding of ACT-R in Constraint Handling Rules (CHR). Therefore analysis techniques from CHR can be used to reason about computational properties of ACT-R models. For example, confluence is the property that a program yields the same result for the same input regardless of the rules that are applied. In ACT-R models, there are often cognitive processes that should always yield the same result while others e.g. implement strategies to solve a problem that could yield different results. In this paper, a decidable confluence criterion for ACT-R is presented. It allows to identify ACT-R rules that are not confluent. Thereby, the modeler can check if his model has the desired behavior.

The sound and complete translation of ACT-R to CHR from prior work is used to come up with a suitable invariant-based confluence criterion from the CHR literature. Proper invariants for translated ACT-R models are identified and proven to be decidable. The presented method coincides with confluence of the original ACT-R models.

Keywords: computational cognitive modeling, confluence, invariants, ACT-R, Constraint Handling Rules

1 Introduction

Computational cognitive modeling is a research field at the interface of cognitive sciences and computer science. It tries to explain human cognition by building detailed computational models of cognitive processes [15]. To support the modeling process, cognitive architectures like Adaptive Control of Thought – Rational (ACT-R) provide the ability to create models of specific cognitive tasks by offering representational formats together with reasoning and learning mechanisms to facilitate modeling [16].

ACT-R is widely employed in the field of computational cognitive modeling. It is defined as a production rule system that offers advanced conflict resolution mechanisms to model learning and competition of different strategies for problem

solving. Therefore, many ACT-R models are highly non-deterministic to resemble the applicability of more than one strategy in many situations. The strategy is chosen depending on information learned from situations in the past.

Confluence is the property of a program that regardless of the order its rules are applied, they finally yield the same result. By identifying the rules that lead to non-confluence, model quality can be improved: It allows to check if the model has the desired behavior regarding competing strategies and e.g. identify rules that interfere with each other unintentionally.

In this paper, we present a decidable confluence test for the abstract operational semantics of ACT-R using confluence analysis tools for CHR. In prior work, we presented a sound and complete embedding of ACT-R in CHR [9,8]. An invariant-based confluence test for CHR [6,11] is used to decide confluence of the translated models with invariants on CHR states that come from the abstract operational semantics of ACT-R. The confluence test identifies the rules that lead to non-confluence supporting the decision if a model has the desired behavior regarding competing strategies.

First the preliminaries are recapitulated in section 2. The main section 3 describes the confluence criterion for ACT-R models. For this purpose, the invariant-based confluence test for CHR is introduced briefly (section 3.1). Then, the ACT-R invariant is defined and a *decidable criterion* for the invariant is given (section 3.2). It is shown that the ACT-R invariant is maintained in the translation. The theoretical foundations to apply the CHR invariant-based confluence test to ACT-R models are derived resulting in a *confluence criterion for terminating ACT-R models* (section 3.3). An example is given in section 3.4.

2 Preliminaries

2.1 Confluence

Confluence is the property of a state transition system that same inputs yield the same results regardless of which rules are applied.

Definition 1 (joinability and confluence [7]). In a state transition system (S, \mapsto) with states S and a transition relation $\mapsto : S \times S$ with reflexive transitive closure \mapsto^* , two states $\sigma_1, \sigma_2 \in S$ are joinable, denoted as $\sigma_1 \downarrow \sigma_2$, if there exists a state σ' such that $\sigma_1 \mapsto^* \sigma'$ and $\sigma_2 \mapsto^* \sigma'$. A state transition system is confluent, if for all states $\sigma, \sigma_1, \sigma_2 : (\sigma \mapsto^* \sigma_1) \land (\sigma \mapsto^* \sigma_2) \rightarrow (\sigma_1 \downarrow \sigma_2)$.

Hence, a program is confluent if for all states that lead to different successor states, those states are joinable. A program is *locally confluent*, if $(\sigma \mapsto \sigma_1) \land (\sigma \mapsto \sigma_2)$ in one transition step and σ_1 and σ_2 are joinable. It can be shown that for all state transition systems local confluence and confluence are equivalent [7]. Figure 1 illustrates (local) confluence.

2.2 Adaptive Control of Thought – Rational (ACT-R)

In this section, ACT-R is introduced briefly. An extensive introduction to the theory can be found in [3,16]. ACT-R is a modular production rule system.



Fig. 1. Confluence and local confluence.

Its data elements are so-called *chunks*. A chunk has a *type* and a set of *slots* (determined by the type) that are connected to other chunks. Hence, human declarative knowledge is represented in ACT-R as a network of chunks. Figure 2 shows an example chunk network that models the representation of an order over natural numbers.



Fig. 2. A chunk network that represents the order of natural numbers 1, 2, 3. The chunks are represented by nodes, the slots by labeled edges. The labels of the nodes are chunk identifiers. Chunks 1, 2 and 3 are of type *number* that has no slots. Chunks *a* and *b* are of type *order* that has a *first* and a *second* slot.

ACT-R's modules are responsible for different cognitive features. For instance, the declarative knowledge (represented as a chunk network) can be found in the *declarative module*. Each module has a set of associated *buffers* that contain at most one chunk. The heart of ACT-R is the *procedural system* that consists of a set of *production rules*. Those rules only have access to the contents of the buffers. They match the contents of the buffer, i.e. they check if the chunks of particular buffers have certain values. If a rule is applicable, it can *modify* particular slots of the chunk in the buffer, *request* the module to put a whole new chunk in its buffer or *clear* a buffer. Modifications and clearings are available directly for the production rule system, whereas requests can take some time while the procedural system is continuing work in parallel.

Syntax of ACT-R We use our simplified syntax in form of first-order terms that can be derived directly from the original syntax [9,8]. The syntax of ACT-R is defined over two disjoint sets of constant symbols C and variable symbols \mathcal{V} . An ACT-R model consists of a set of types \mathbb{T} with type definitions and a set of rules Σ .

A production rule has the form $\mathcal{L} \Rightarrow R$ where \mathcal{L} is a finite set of buffer tests. A buffer test is a first-order term of the form =(b, t, P) where the buffer $b \in C$, the type $t \in C$ and $P \subseteq C \times (C \cup V)$ is a set of slot-value pairs (s, v) where $s \in C$ and $v \in C \cup V$. This means that only the values in the slot-value pairs can consist of both constants and variables.

The right-hand side $R \subseteq \mathcal{A}$ of a rule is a finite set of actions where $\mathcal{A} = \{a(b,t,P) \mid a \in A, b \in \mathcal{C}, t \in \mathcal{C} \text{ and } P \subseteq \mathcal{C} \times (\mathcal{C} \cup \mathcal{V})\}$. Hence, an action is a term of the form a(b,t,P) where the functor a of the action is in A, the set of action symbols, the first argument b is a constant (denoting a buffer), the second argument is a constant t denoting a type, and the last argument is a set of slot-value pairs, i.e. a pair of a constant and a constant or variable. Usually, the action symbols are defined as $A := \{=, +, -\}$ for modifications, requests and clearings respectively. Only one action per buffer is allowed, i.e. if $a(b,t,P) \in R$ and $a'(b',t',P') \in R$, then $b \neq b'$ [4].

We assume the rules to be in so-called *set normal form* that requires the slot tests of a rule to be total and unique with respect to the type of the test. This means that each slot defined by the type of the tested chunk must appear at most once in the set of slot-value pairs. Every rule can be transformed to set normal form [8].

Operational Semantics of ACT-R For the understanding of this paper, it is sufficient to define ACT-R states and rules formally. The formal definition of the operational semantics can be found in [9,8]. We define the operational semantics of ACT-R through our CHR translation that has been first presented in [9] and in its most current form in [8]. Since the translation is sound and complete, we omit the formal definition of the ACT-R semantics here, since it would only distract from the contribution of this paper.

Definition 2 (chunk types, chunk stores). A typing function $\tau : \mathbb{T} \to 2^{\mathcal{C}}$ maps each type from the set $\mathbb{T} \subseteq \mathcal{C}$ to a finite set of allowed slot names. A chunk store Δ is a multi-set of tuples (t, val) where $t \in \mathbb{T}$ is a chunk type and $val : \tau(t) \to \Delta$ is a function that maps each slot of the chunk (determined by the type t) to another chunk. Each chunk store Δ has a bijective identifier function $id_{\Delta} : \Delta \to \mathcal{C}$ that maps each chunk of the multi-set a unique identifier.

Additional information represents the inner state of the modules and so-called sub-symbolic information used in ACT-R implementations to model cognitive features like forgetting, latencies and conflict resolution. The information is expressed as a conjunction of predicates from first-order logic. We now define ACT-R states as follows:

Definition 3 (cognitive state, ACT-R state). A cognitive state γ is a function $\mathbb{B} \to \Delta \times \mathbb{R}_0^+$ that maps each buffer to a chunk and a delay. The delay decides at which point in time the chunk in the buffer is available to the production system. A delay d > 0 indicates that the chunk is not yet available to the production system. This implements delays of the processing of requests.

An ACT-R state is a tuple $\langle \Delta; \gamma; v \rangle$ where γ is a cognitive state and v is a multi-set of ground, atomic first order predicates (called additional information).

2.3 Constraint Handling Rules (CHR)

In this section, syntax and semantics of CHR are summarized briefly. For an extensive introduction to CHR, its semantics, analysis and applications, we refer

to [7]. We use the latest definition of the state transition system of CHR that is based on state equivalence [12]. The definitions from those canonical sources are now reproduced.

The syntax of CHR is defined over a set of variables, a set of function symbols with arities and a set of predicate symbols with arities that is disjointly composed of *CHR constraint symbols* and *built-in constraint symbols*. The set of constraint symbols contains at least the symbols = /2, $\top/0$ and $\perp/0$. In this paper, we allow the terms to be sets of terms as they can be simply represented as lists in implementations. For a constraint symbol c/n and terms t_1, \ldots, t_n over the variables and function symbols, $c(t_1, \ldots, t_n)$ is called a *CHR constraint* or a *built-in constraint*, depending on the constraint symbol. We now define the notion of CHR states.

Definition 4 (CHR state). A CHR state is a tuple $\langle \mathbb{G}; \mathbb{C}; \mathbb{V} \rangle$ where the goal \mathbb{G} is a multi-set of constraints, the built-in constraint store \mathbb{C} is a conjunction of built-in constraints and \mathbb{V} is a set of global variables.

All variables occurring in a state that are not global are called local variables.

CHR states can be modified by rules that together form a CHR program. For the sake of brevity, we only consider simplification rules, as they are the only type of rules needed for the understanding of the paper.

Definition 5 (CHR program). A CHR program is a finite set of rules of the form $r @ H \Leftrightarrow G | B_c, B_b$ where r is an optional rule name, the heads H are multi-sets of CHR constraints, the guard G is a conjunction of built-in constraints and the body is a multi-set of CHR constraints B_c and a conjunction of built-in constraints B_b . If G is empty, it is interpreted as the built-in constraint \top .

Informally, a rule is applicable, if the head matches constraints from the store \mathbb{G} and the guard holds, i.e. is a consequence of the built-in constraints \mathbb{C} . In that case, the constraints matching H are removed and the constraints from B_c , B_b and G are added.

In the context of the operational semantics, we assume a constraint theory \mathcal{CT} for the interpretation of the built-in constraints. We define an equivalence relation over CHR states.

Definition 6 (CHR state equivalence [11,12]). Let $\rho := \langle \mathbb{G}; \mathbb{C}; \mathbb{V} \rangle$ and $\rho' := \langle \mathbb{G}'; \mathbb{C}'; \mathbb{V}' \rangle$ be CHR states with local variables \bar{y}, \bar{y}' that have been renamed apart. $\rho \equiv \rho'$ if and only if

$$\mathcal{CT} \models \forall (\mathbb{C} \to \exists \bar{y}'. ((\mathbb{G} = \mathbb{G}') \land \mathbb{C}')) \land \forall (\mathbb{C}' \to \exists \bar{y}. ((\mathbb{G} = \mathbb{G}') \land \mathbb{C}))$$

where $\forall F$ denotes the universal closure of formula F.

The operational semantics is now defined by the following transition scheme over equivalence classes of CHR states i.e. $[\rho] := \{\rho' \mid \rho' \equiv \rho\}$

Definition 7 (operational semantics of CHR [11,12]). For a CHR program the state transition system over CHR states and the rule transition relation \mapsto is

defined as the following transition scheme:

$$\frac{r @ H \Leftrightarrow G \mid B_c, B_b}{[\langle H \uplus \mathbb{G}; G \land \mathbb{C}; \mathbb{V} \rangle] \mapsto^r [\langle B_c \uplus \mathbb{G}; G \land B_b \land \mathbb{C}; \mathbb{V} \rangle]}$$

Thereby, r is a variant of a rule in the program such that its local variables are disjoint from the variables occurring in the representative of the pre-transition state. We may just write \mapsto instead of \mapsto^r if the rule r is clear from the context.

2.4 Translation of ACT-R to CHR

We briefly summarize the translation of ACT-R models to CHR first presented in [9,8]. Since the translation is proven to be sound and complete [8], we explain the operational semantics of ACT-R with the help of the translation.

Definition 8 (translation of abstract states). An abstract ACT-R state $\sigma := \langle \Delta; \gamma; v \rangle$ can be translated to the following CHR state:

Thereby, [val] denotes the explicit relational notation of the function val as a set of tuples. We denote the translation of an ACT-R state σ by $chr(\sigma)$.

The chunk store is represented by a delta constraint that contains a set of chunk/3 terms representing the chunks with their identifiers, types and slot-value pairs.

For every buffer of the given architecture, there is a constraint *gamma* with buffer name, chunk identifier and delay. Since γ is a total function, every buffer has exactly one *gamma* constraint. Additional information is represented directly as built-in constraints.

Definition 9 (translation of rules). Let $cogstate(\mathbb{B}) := \{(b, C_b) \mid b \in \mathbb{B}\}$ be the relation that connects each buffer with a variable C_b . An ACT-R rule in set-normal form $r := \mathcal{L} \Rightarrow \mathcal{R}$ can be translated to a CHR rule of the form:

 $r @ delta(D) \uplus \{gamma(b, C_b, E_b) \mid b \in \mathbb{B}\}$

 \Leftrightarrow

$$\bigwedge \quad (chunk(C_b, t, P) \text{ in } D \land E_b = 0) \mid$$

 $\textbf{=}(b,t,P){\in}\mathcal{L}$

 $\{ delta(D^*) \} \uplus \{ gamma(b, C_b^{**}, resdelay(b)) \mid b \in \mathbb{B} \land a(b, t, P) \in R \} \\ \uplus \{ gamma(b, C_b, E_b) \mid b \in \mathbb{B} \land a(b, t, P) \notin R \},$

 $\bigwedge_{\alpha=a(b,t,P)\in\mathcal{R}} action(\alpha,D,cogstate(\mathbb{B}),D_b^*,C_b^*,E_b^*)$

 $\land \ merge([D_b^*:a(b,t,P)\in R],D')\} \land merge([D,D'],D^*) \ \uplus$

$$\wedge \bigwedge_{a(b,t,P)\in R} map(D,D',C_b^*,C_b^{**}).$$

Note that ACT-R constants and variables from C and V are implicitly translated to corresponding CHR variables.

We denote the translation of a rule r by chr(r) and the translation of an ACT-R model Σ that is a set of ACT-R rules by $chr(\Sigma)$. Thereby, $chr(\Sigma) := \{chr(r) \mid r \in \Sigma\}$.

The rule removes the *delta* and all gamma constraints from the store. It binds the translation of the chunk store Δ to the variable D. For all buffers b, each variable C_b is bound to the chunk identifier of the chunk in b, i.e. $C_b = id_{\Delta}(\gamma(b))$. The guard now performs all buffer tests =(b, t, P) from the ACT-R rule by testing if a chunk term $chunk(C_b, t, P)$ is in the translated chunk store D that has type t and matches all slot-value pairs in P. The ACT-R variables in P are bound to the values in the state.

In the body, the built-in constraints *action* perform the actions of the ACT-R rule as defined by the architecture. An *action* constraint gets the action term α of the rule (with all variables bound through the matching), the original chunk store and a representation of the cognitive state. Since the C_b have been bound in the matching, it consists of tuples that connect each buffer b with the chunk identifier it holds.

The action built-in constraint returns a chunk store D_b^* , a chunk identifier C_b^* that represents the resulting chunk from the request and a result delay E_b^* . The merge constraints merge the chunk stores of all actions with the original store D to the store D^* . The result of merging two chunk stores can vary from implementation to implementation, but has to obey some rules defined in [8]. One can think of it as a multi-set union. As chunk identifiers might change in the merging process, the built-in map maps the chunk identifier of the results to the corresponding identifiers in the merged store.

Then, a new Δ constraint with the resulting chunk store D^* is added as well as the gamma constraints. If the buffer b has been part of an action, then it is altered such that it holds the resulting chunk identifier C_b^{**} after the merge and the resulting delay E_b^* . If it was not part of an action, its parameters C_b (the chunk identifier) and E_b (the delay) remain unchanged. This is possible, since the chunk merging guarantees that chunks in the original chunk store D the constraint gamma is referring to, are also part of the merged chunk store D^* .

Example 1 (counting). We now give an example ACT-R rule to explain its operational semantics. A classical example in ACT-R is counting by recalling count order facts. The model uses chunks of type *order* as illustrated in figure 2. An *order* chunk has a *first* and a *second* slot that link two chunks representing natural numbers in the right order. Additionally, we define a second chunk type g that memorizes the current number in the counting process. The main rule is defined as:

 $=(goal, g, \{(current, X)\}), =(retrieval, order, \{(first, X), (second, Y)\})$ $\Rightarrow =(goal, g, \{(current, Y)\}), +(retrieval, order, \{(first, Y)\})$

The left-hand side tests if there is a chunk of type g in the goal buffer. The value of its *current* slot is bound to variable X by the matching. The second buffer test checks the *retrieval* buffer for a chunk of type *order* that has X in its *first* slot. The value of the *second* slot is bound to variable Y.

The right-hand side modifies the chunk in the goal buffer such that Y is written to the *current* slot. The second action requests the *retrieval* buffer for an *order* chunk that has Y in its *first* slot. As soon as the requested chunk is available, the program can apply the rule again. The head and guard of the CHR translation $H \Leftrightarrow G \mid B$ of the rule is

 $\begin{aligned} H &:= \{ delta(D), gamma(g, C_g, 0), gamma(retrieval, C_r, 0) \}, \\ G &:= chunk(C_g, g, \{ (current, X) \}) \text{ in } D \land \\ chunk(C_r, order, \{ (first, X), (second, Y) \}) \text{ in } D. \end{aligned}$

3 Confluence Criterion for ACT-R

This section is the main contribution of the paper. We gradually develop a decidable criterion for confluence of ACT-R using the CHR embedding.

Therefor, a brief introduction to invariant-based confluence analysis for CHR is given that extends the standard confluence criterion to handle invariants that must hold for the regarded states. We then define the ACT-R invariant \mathcal{A} that is satisfied if a CHR state has been derived from an ACT-R state. Then a decidable criterion for the invariant is presented and it is shown that the invariant is maintained in translated ACT-R models. It is shown how invariant-based confluence analysis for CHR can be applied to decide ACT-R confluence.

3.1 Invariant-based Confluence

We now give a brief introduction to invariant-based confluence analysis for CHR. The first results stem from [6]. We summarize the main theorem of the improved version that can be found in [11, section 14].

The main idea of the confluence criterion is that heads and guards of the rules are overlapped to an overlap state. Then both overlapping rules are applied to this state forming a critical pair that is checked for joinability for all possible overlap states. An overlap is defined as follows:

Definition 10 (overlap and critical pairs [7,11]). For any two (not necessarily different) rules of a CHR program with renamed apart variables of the form $r @ H \Leftrightarrow G | B_c, B_b \text{ and } r' @ H' \Leftrightarrow G' | B'_c, B'_b, \text{ let } O \subseteq H, O' \subseteq H' \text{ such that}$ for $B := (O = O') \land G \land G'$ it holds that $C\mathcal{T} \models \exists.B \text{ and } O \neq \emptyset$, then the state

$$\sigma = \langle R \uplus R' \uplus O; B; \mathbb{V} \rangle$$

is called an overlap of r and r' where $R := H \setminus O$, $R' := H' \setminus O'$ and \mathbb{V} is the set of all variables occurring in heads and guards of both rules. The pair of states $\sigma_1 := \langle R' \uplus B_c; B \wedge B_b; \mathbb{V} \rangle$ and $\sigma_2 := \langle R \uplus B'_c; B \wedge B'_b; \mathbb{V} \rangle$ is a critical pair of the overlap σ .

CHR has the monotonicity property. It states that all rules that are applicable in a state, are also applicable in any larger state. This idea can be exploited to reason from joinable overlap states about local confluence and therefore confluence of a CHR program. The problem with invariant-based confluence is that the idea of using monotonicity to reason about larger states does not work for states where the invariant does not hold. An overlap that does not satisfy the invariant makes all information about this state irrelevant [11, p. 79]. The idea of the invariant-based confluence theorem for CHR is to extend all states where the invariant does not hold such that the invariant is repaired and include the extended states in the confluence test. Since in general there are infinitely many extensions that maintain the invariant, only minimal extensions according to a partial order defined in [11] have to be considered. Then, monotonicity can be applied again.

Theorem 1 (invariant-based confluence for CHR [11]). For an invariant \mathcal{I} , let $\Sigma^{\mathcal{I}}([\rho]) := \{[\rho'] \mid [\rho'] \text{ is an extension of } [\rho] \text{ such that } \mathcal{I} \text{ holds } \}$ be the set of satisfying extensions of $[\rho]$. The set $\mathcal{M}^{\mathcal{I}}([\rho])$ is the set of minimal elements of $\Sigma^{\mathcal{I}}([\rho])$ w.r.t. the partial order on states defined in [11].

Let \mathcal{P} be a CHR program and $\mathcal{M}^{\mathcal{I}}([\rho])$ be well-defined for all overlaps ρ . \mathcal{P} is locally confluent with respect to \mathcal{I} if and only if for all overlaps ρ with critical pairs (ρ_1, ρ_2) and all $[\rho_m] \in \mathcal{M}^{\mathcal{I}}([\rho])$ holds that $[\rho_1]$ extended by $[\rho_m]$ and $[\rho_2]$ extended by $[\rho_m]$ are joinable. We then say that \mathcal{P} is \mathcal{I} -(locally) confluent.

There are two problems with this result making it possibly undecidable: The invariant could be undecidable and the set of minimal elements can be infinitely large. We will show that in the case of the ACT-R invariant that we use for our confluence test, the set of satisfying extensions is empty and the invariant is decidable. Hence, it is not necessary for the understanding of this paper how the partial order on states and therefore the set of minimal elements is defined formally, since the set of satisfying extensions is already empty for the ACT-R invariant. The ACT-R invariant is defined in the following section.

3.2 ACT-R Invariant

To reason about confluence of ACT-R models in CHR, we need an invariant that restricts the CHR state space to states that stem from a valid ACT-R state. In the following example, we show how overlapping translated ACT-R rules can lead to overlap states that do not describe a valid ACT-R state.

Example 2. Let $\{delta(D), gamma(B, C, 0)\} \Leftrightarrow chunk(C, T, P) \text{ in } D \mid \dots$ be a CHR rule that has been obtained from an ACT-R rule. By overlapping the rule with itself, we could get

$$\sigma := \langle delta(D), gamma(B, C, 0), gamma(B, C', 0); \\ chunk(C, T, P) \text{ in } D \land chunk(C', T', P') \text{ in } D; \mathbb{V} \rangle.$$

However, this state does not stem from a valid ACT-R state, since γ is a function with only one value for each buffer and therefore the translation of an ACT-R state can never contain two gamma constraints for the same buffer B.

In the following, we define the ACT-R invariant \mathcal{A} on CHR states that limits the state space to states that stem from valid ACT-R states. We show that the invariant is decidable by breaking it down to five fine grained invariants. We also show that it actually defines an invariant for translated ACT-R models.

Definition 11 (ACT-R invariant). Let $[\rho]$ be a CHR state. The ACT-R invariant \mathcal{A} holds if and only if there is an ACT-R state σ such that $\rho \equiv chr(\sigma)$.

Basically, this means that $\mathcal{A}([\rho])$ holds if $[\rho]$ is the valid translation of an ACT-R state. However, by this definition it is hard to decide if a CHR state satisfies the invariant.

We now show some decidable sub-invariants on CHR states and prove that their conjunction is equivalent to \mathcal{A} . For this purpose, we define an auxiliary function *ids* that returns the set of chunk identifiers for a set of *chunk/3* terms.

Definition 12 (chunk identifiers). Let d be a set. Then

$$ids(d) := \{c \mid chunk(c,t,p) \in d\}$$

is the set of chunk identifiers of the set d.

The sub-invariants mainly consist of *uniqueness* invariants, i.e. they require that there is only one constraint of a certain kind for a class of arguments, and *functional dependency* invariants, i.e. that certain sets that represent relations appearing in constraints are functions. Eventually, the constraints that can be be used in a state are restricted.

Theorem 2 (ACT-R invariants). Let $\rho \equiv \langle \mathbb{G}; \mathbb{C}; \mathbb{V} \rangle$ be a CHR state. We define the following sub-invariants:

1. unique chunk store

 $\mathcal{A}_1([\rho]) \leftrightarrow \text{ There is exactly one constraint } delta(d) \in \mathbb{G} \text{ for some ground set} \\ d. \text{ For all elements } e \in d, \text{ it holds that there exist } c \in \mathcal{C}, t \in \mathbb{T}, p \in \mathcal{C} \times \mathcal{C}, s \in \\ \tau(t), v \in \mathcal{C} \text{ such that } e = chunk(c, t, p) \text{ and } p = \{(s, v) \mid s \in \tau(t) \land v \in ids(d)\}.$

- 2. functional dependency of cognitive state $\mathcal{A}_2([\rho]) \leftrightarrow$ For all buffers $b \in \mathbb{B}$ there is exactly one gamma $(b, c, e) \in \mathbb{G}$ where $c \in ids(d)$ for some delta $(d) \in \mathbb{G}$ and $e \in \mathbb{R}^+_0$.
- 3. unique chunk identifiers $\mathcal{A}_3([\rho]) \leftrightarrow$ For all chunk identifiers $c \in \mathcal{C}$ and constraints $delta(d) \in \mathbb{G}$, if $chunk(c,t,p) \in d$, then there is no other term $chunk(c,t',p') \in d$.
- 4. functional dependency of slot-value pairs $\mathcal{A}_4([\rho]) \leftrightarrow$ For all constraints $delta(d) \in \mathbb{G}$, terms chunk(c,t,p) in set d and (s,v) in set p, there is no other term (s,v') in p.
- 5. allowed constraints

 $\mathcal{A}_5([\rho]) \leftrightarrow \text{In } \mathbb{G}$ there are only delta/1 and gamma/3 constraints, only syntactic equality = /2 and the allowed constraints defined by the ACT-R architectures appear in \mathbb{C} and $[\rho]$ is ground.

For all CHR states $[\rho]$ it holds that $\mathcal{A}([\rho]) \leftrightarrow \bigwedge_{i=1}^{5} \mathcal{A}_{i}([\rho])$.

- *Proof.* if direction If $\mathcal{A}([\rho])$, then $[\rho]$ is the product of the translation of an ACT-R state. It follows directly from definition 8 that in that case, $\mathcal{A}_1([\rho])$, $\mathcal{A}_2([\rho])$, $\mathcal{A}_3([\rho])$, $\mathcal{A}_4([\rho])$ and $\mathcal{A}_5([\rho])$ hold.
- only-if direction We have to show that for all CHR states $[\rho]$ where the invariants $\mathcal{A}_1([\rho])$, $\mathcal{A}_2([\rho])$, $\mathcal{A}_3([\rho])$, $\mathcal{A}_4([\rho])$ and $\mathcal{A}_5([\rho])$ hold, there is an ACT-R state σ such that $\rho \equiv chr(\sigma)$. Let $[\rho] := [\langle \mathbb{G}; \mathbb{C}; \mathbb{V} \rangle]$.

We construct the ACT-R state $\sigma := \langle \Delta; \gamma; v \rangle$. Since $\mathcal{A}_1([\rho])$, there is exactly one delta(d) constraint for a set d and all elements in d are of the form chunk(c,t,p) where $c \in \mathcal{C}, t \in \mathbb{T}$ and p is a set of elements (s,v) with $s \in \tau(t)$ and $v \in ids(d)$. The set p is total with respect to s and the v are chunk identifiers that appear in d. Due to \mathcal{A}_4 , there is exactly one $(s,v) \in p$ for each $s \in \tau(t)$, hence p is the relational representation of a value function. The invariant \mathcal{A}_3 guarantees that the chunk identifiers are unique.

We define $\Delta := \{(t, p) \mid chunk(c, t, p) \in d\}$ with the identifier function $id_{\Delta} := \{((t, p), c) \mid chunk(c, t, p)\}.$

Due to invariant \mathcal{A}_2 , the cognitive state can then be defined for all $b \in \mathbb{B}$ such that $\gamma(b) := (id_{\Lambda}^{-1}(c), e)$ for each $gamma(b, c, e) \in \mathbb{G}$.

Since $\mathcal{A}_5([\rho])$, $[\rho]$ is ground. Hence, we can find another representative of the state with $\rho \equiv \langle \mathbb{G}'; \mathbb{C}'; \emptyset \rangle$, that applies all equality constraints X=t in \mathbb{C} such that only constants appear in \mathbb{G}' and \mathbb{C}' and \mathbb{C}' only consists of allowed predicates defined by the ACT-R architecture. Therefore, we can set $v := \mathbb{C}'$. From the construction of σ it is clear that $\rho \equiv chr(\sigma)$.

The invariants $\mathcal{A}_1, \ldots, \mathcal{A}_5$ are obviously decidable. Since they are equivalent to the ACT-R invariant \mathcal{A} , theorem 2 gives us a decidable criterion for the ACT-R invariant \mathcal{A} .

In the next step, we show that the ACT-R invariant \mathcal{A} is maintained by transitions that come from a translated ACT-R program, i.e. that it really is an invariant.

Lemma 1. Let \mapsto be the state transition relation derived from the translation of an ACT-R model and $[\rho]$ a CHR state with $\mathcal{A}([\rho])$. If $[\rho] \mapsto [\rho']$, then $\mathcal{A}([\rho'])$.

Proof. We are going to use soundness and completeness [8] to prove this.

Let $[\rho]$ be a CHR state with $\mathcal{A}([\rho])$. Since $\mathcal{A}([\rho])$, there is an ACT-R state σ with $\rho \equiv chr(\sigma)$. Due to the sound and complete embedding of ACT-R in CHR, there is an ACT-R state σ' with $\rho' \equiv chr(\sigma')$. Hence, $\mathcal{A}([\rho'])$ holds.

3.3 Invariant-Based Confluence Test

We want to use theorem 1 [11, p. 83, theorem 6] to prove confluence of all states $[\rho]$ that satisfy the ACT-R invariant, i.e. where $\mathcal{A}([\rho])$. Therefore, we have to construct the set $\Sigma^{\mathcal{A}}([\rho])$ for each state $[\rho]$ that does not satisfy \mathcal{A} . It contains all states that can be merged to $[\rho]$ such that they satisfy \mathcal{A} (see theorem 1). The minimal elements in this set have to be considered in the confluence test.

We will see that for all states $[\rho]$ that do not satisfy \mathcal{A} , the set of minimal elements is empty. Intuitively, this means that there are no states that can extend $[\rho]$ such that it satisfies \mathcal{A} .

Lemma 2 (minimal elements for \mathcal{A}). Let \mathcal{A} be the ACT-R invariant as defined in definition 2. For all states $[\rho]$ such that $\mathcal{A}([\rho])$ does not hold, $\Sigma^{\mathcal{A}}([\rho]) = \emptyset$ and therefore $\mathcal{M}^{\mathcal{A}}([\rho]) = \emptyset$.

Proof. Let $[\rho] := [\langle \mathbb{G}; \mathbb{C}; \mathbb{V} \rangle]$. We use theorem 2 that allows us to analyze the individual sub-invariants:

- 1. If \mathcal{A}_1 is violated, there are the following cases:
 - There are two constraints delta(d), $delta(d') \in \mathbb{G}$. We cannot extend $[\rho]$ (i.e. add constraints) to satisfy \mathcal{A}_1 .
 - There is only one unique $delta(d) \in \mathbb{G}$, with elements that do not have the required form. Again, no constraints can be added to satisfy \mathcal{A}_1 .
- 2. If \mathcal{A}_2 is violated, there are two constraints $gamma(b, c, e), gamma(b', c', e') \in \mathbb{G}$. We cannot satisfy \mathcal{A}_2 for such a state.
- 3. The proof is analogous for \mathcal{A}_3 and \mathcal{A}_4 .
- 4. If \mathcal{A}_5 is violated, there are other constraints then *delta* or *gamma* in \mathbb{G} or other than the allowed constraints defined by the architecture in \mathbb{C} . This cannot be repaired by extending \mathbb{G} or \mathbb{C} .

We can directly apply theorem 1: For all overlaps ρ where $\mathcal{A}([\rho])$ holds, the set of minimal elements is $\mathcal{M}^{\mathcal{A}}([\rho]) = \{[\rho_{\emptyset}]\}$ [11, p.80, lemma 13.13] where $\rho_{\emptyset} := \langle \emptyset; \top; \emptyset \rangle$ is the empty CHR state. Hence, for overlaps where \mathcal{A} holds, we only have to show joinability of the critical pairs that stem from the overlap itself. This coincides with the regular confluence test of CHR as defined in [7].

For all overlaps ρ where $\mathcal{A}([\rho])$ does not hold, the set of minimal elements is $\mathcal{M}^{\mathcal{A}}([\rho]) = \emptyset$ by lemma 2. Therefore, no critical pairs have to be tested. We summarize this in the following theorem.

Theorem 3 (A-local confluence). A CHR program is A-local confluent if and only if for all critical pairs (ρ_1, ρ_2) with overlap ρ for which $\mathcal{A}(\rho)$, it is $\rho_1 \downarrow \rho_2$.

Proof. This follows directly from theorem 1 and lemma 2 for overlaps where $\mathcal{A}([\rho])$ does not hold. For overlaps with $\mathcal{A}([\rho])$, the unique minimal element is the empty state $[\rho_{\emptyset}] := [\langle \emptyset; \top; \emptyset \rangle]$ which is the neutral element for state merging [11, lemma 13.13, p. 80]. Therefore, if $\mathcal{A}([\rho])$ holds, it suffices to test the critical pairs that stem from $[\rho]$ by theorem 1.

We now have a criterion to decide \mathcal{A} -confluence of \mathcal{A} -terminating CHR programs that have been translated from an ACT-R model. In the next theorem, we show that \mathcal{A} -confluence of such CHR programs coincides with ACT-R confluence. Therefore, the confluence criterion is applicable to decide confluence of ACT-R models.

Theorem 4 (confluence in ACT-R). Let M be an ACT-R model. Then M is terminating and confluent if and only if chr(M) is A-terminating and A-confluent.

13

Proof. A-termination is maintained through soundness and completeness. We now show that confluence for terminating models and their CHR counterparts coincides. Confluence is defined as $(\sigma \mapsto^* \sigma_1) \land (\sigma \mapsto^* \sigma_2) \rightarrow (\sigma_1 \downarrow \sigma_2)$ for all states $\sigma, \sigma_1, \sigma_2$. It remains to show that joinability in ACT-R and CHR are equivalent, i.e. $(\sigma_1 \downarrow \sigma_2) \leftrightarrow ([chr(\sigma_1)] \downarrow [chr(\sigma_2)])$.

- **If-direction** If $(\sigma_1 \downarrow \sigma_2)$, there is a state σ' such that $\sigma_1 \rightarrow^* \sigma'$ and $\sigma_2 \rightarrow^* \sigma'$. Due to soundness and completeness of the embedding, we have that $[chr(\sigma_1)] \rightarrow^* [chr(\sigma')]$ and $[chr(\sigma_2)] \rightarrow^* [chr(\sigma')]$.
- **Only-if-direction** This is analogous. We just have to construct the ACT-R state from the joined CHR state $[\rho']$. Since $\mathcal{A}([\rho'])$ holds by lemma 1, this state exists.

3.4 Example: Counting

We continue our example 1. We assume that each number chunk only appears in at most one *order* chunk at *first* or *second* position. This means that the model has learned a stable order on the numbers and hence requests to the declarative module are deterministic. It is clear that this example model terminates for finite declarative memories. Therefore, we can apply our confluence criterion.

The rule can overlap with itself, e.g. $\langle delta(D), delta(D'), \ldots; \ldots; \ldots \rangle$. This state invalidates invariant \mathcal{A}_1 and hence is not part of the confluence test. Another overlap is $\langle delta(D), gamma(g, C_g, 0), gamma(g, C'_g, 0), \ldots; \ldots; \ldots \rangle$. It violates invariant \mathcal{A}_2 , because it has two gamma constraints for the same buffer.

All overlaps consist of the following built-in store:

$$\langle delta(D), \ldots; chunk(C_g, g, \{(current, X)\}) \ in \ D \\ \wedge \ chunk(C_g, g, \{(current, X')\}) \ in \ D, \ldots; \{D, X, X', \ldots\} \rangle.$$

By invariant \mathcal{A}_3 it must be X = X', because otherwise there were two different *chunk* terms in the same chunk store with the same chunk identifier.

The overlap $\langle H; G; \mathbb{V} \rangle$ that only consists of the head and guard of the rule where \mathbb{V} contains all variables of H and G is joinable, because we assumed determinism of requests, i.e. there is only one possible result chunk for each request. It can be seen that all possible overlaps in this small example invalidate the ACT-R invariant \mathcal{A} or are joinable. Therefore, the model consisting only of this one counting rule is confluent. If we would assume an agent that has not learned a stable order of numbers, yet, i.e. there are numbers with different successors, the model would not be confluent. The confluence test constructs minimal representations of the states that are not joinable, i.e. giving an insight to the reason why a model is not confluent. This allows to decide whether the model has the desired behavior when it comes to different available strategies.

4 Related Work

There exist CHR embeddings of other rule-based approaches. The results on invariant-based confluence analysis have been used successfully to the embedding of graph transformation systems in CHR [10,13].

In the context of ACT-R, there are – to the best of our knowledge – no other approaches that deal with confluence so far. There have been other approaches to formalize the architecture with the aim to reason about cognitive models. For instance, F-ACT-R [2,1] formalizes the architecture of ACT-R to simplify comparison of different models or to use model checking techniques. In [14] mathematical reformulations of ACT-R models are used for parameter optimization by mathematical optimization techniques.

5 Conclusion

In this paper, we have shown a decidable confluence test for the abstract operational semantics of ACT-R. A confluence test can help to improve ACT-R models by identifying the rules that inhibit confluence. This enables the modeler to decide about the correct behavior of the model regarding competing strategies. In our approach, we use the sound and complete embedding of ACT-R in CHR to apply the invariant-based confluence criterion for CHR to reason about ACT-R confluence, since standard CHR confluence is too strict.

We have defined the ACT-R invariant \mathcal{A} on CHR states such that it is satisfied for all states that stem from a valid ACT-R state. The first main result is a decidable criterion for the ACT-R invariant (theorem 2).

Furthermore, the theoretical foundations for applicability of CHR invariantbased confluence for the ACT-R invariant \mathcal{A} are established. This leads to the second main result: an invariant-based CHR \mathcal{A} -confluence test (theorem 3).

Eventually, it is shown that \mathcal{A} -confluence coincides with ACT-R confluence (theorem 4). This makes our CHR approach applicable to decide ACT-R confluence. The criterion is decidable as long as the constraint theories behind the actions are decidable, because the invariant is decidable and the preconditions for the invariant-based confluence test are satisfied in the context of ACT-R.

For the future, we want to investigate how the approach can be extended to confluence modulo equivalence [5], since ACT-R confluence can be too strict due to possibly differing chunk identifiers in the processing of the production rules. An equivalence relation on chunk networks that is defined as a special form of graph isomorphism could abstract from chunk identifiers making a chunk store more declarative. By summarizing possible outcomes of a model in equivalence classes, confluence modulo equivalence can also help to reason about correctness of a model. Confluence modulo this equivalence relation would then guarantee that the model always gives a result of a certain kind defined by the equivalence class. For instance, it would be possible to check if a model always yields a chunk of a certain type, e.g. a number or an order chunk.

Reasoning about requests to modules that appear in a confluence proof can be extended by specific constraint theories on the modules that integrate domainspecific knowledge about the model. This idea can be extended by allowing for model-specific constraint theories. For instance, the integration of domain-specific knowledge on chunk types in the context of a particular cognitive model could improve reasoning about module requests in such models.

References

- Albrecht, R., Westphal, B.: Analysing psychological theories with F-ACT-R. In: Proceedings of the 12th Biannual conference of the German cognitive science society (Gesellschaft für Kognitionswissenschaft). Cognitive Processing, vol. 15 (Suppl. 1), pp. 27–28. Springer (2014)
- Albrecht, R., Westphal, B.: F-ACT-R: defining the ACT-R architectural space. In: Proceedings of the 12th Biannual conference of the German cognitive science society (Gesellschaft für Kognitionswissenschaft). Cognitive Processing, vol. 15 (Suppl. 1), pp. 79–81. Springer (2014)
- Anderson, J.R., Bothell, D., Byrne, M.D., Douglass, S., Lebiere, C., Qin, Y.: An integrated theory of the mind. Psychological Review 111(4), 1036–1060 (2004)
- 4. Bothell, D.: ACT-R 6.0 Reference Manual Working Draft. Department of Psychology, Carnegie Mellon University, Pittsburgh, PA
- Christiansen, H., Kirkeby, M.H.: On proving confluence modulo equivalence for Constraint Handling Rules. Formal Aspects of Computing 29(1), 57–95 (2017), http://dx.doi.org/10.1007/s00165-016-0396-9
- Duck, G.J., Stuckey, P.J., Sulzmann, M.: Observable confluence for Constraint Handling Rules. In: Logic Programming, pp. 224–239. Springer (2007)
- 7. Frühwirth, T.: Constraint Handling Rules. Cambridge University Press (2009)
- 8. Gall, D., Frühwirth, T.: An Operational Semantics for the Cognitive Architecture ACT-R and its Translation to Constraint Handling Rules. ArXiv e-prints (Feb 2017)
- Gall, D., Frühwirth, T.: Translation of Cognitive Models from ACT-R to Constraint Handling Rules. In: Alferes, J.J., Bertossi, L., Governatori, G., Fodor, P., Roman, D. (eds.) Rule Technologies. Research, Tools, and Applications: 10th International Symposium, RuleML 2016, Stony Brook, NY, USA, July 6-9, 2016. Proceedings. pp. 223–237. Springer International Publishing (2016), http://dx.doi.org/10.1007/ 978-3-319-42019-6 15
- Raiser, F.: Graph transformation systems in CHR. In: Dahl, V., Niemelä, I. (eds.) Logic Programming: 23rd International Conference, ICLP 2007, Porto, Portugal, September 8-13, 2007. Proceedings. pp. 240–254. Springer Berlin Heidelberg, Berlin, Heidelberg (2007), http://dx.doi.org/10.1007/978-3-540-74610-2_17
- 11. Raiser, F.: Graph Transformation Systems in Constraint Handling Rules: Improved Methods for Program Analysis. Ph.D. thesis, Ulm University, Germany (2010), http://dx.doi.org/10.18725/OPARU-1742
- Raiser, F., Betz, H., Frühwirth, T.: Equivalence of CHR states revisited. In: Raiser, F., Sneyers, J. (eds.) 6th International Workshop on Constraint Handling Rules (CHR). pp. 33–48. KULCW, Technical report CW 555 (July 2009)
- Raiser, F., Frühwirth, T.: Analysing graph transformation systems through Constraint Handling Rules. Theory Practice of Logic Programming 11(1), 65–109 (Jan 2011)
- Said, N., Engelhart, M., Kirches, C., Körkel, S., Holt, D.V.: Applying mathematical optimization methods to an ACT-R instance-based learning model. PloS one 11(7), e0158832 (2016)
- Sun, R.: Introduction to computational cognitive modeling. In: Sun, R. (ed.) The Cambridge Handbook of Computational Psychology, pp. 3–19. Cambridge University Press, New York (2008)
- Taatgen, N.A., Lebiere, C., Anderson, J.: Modeling paradigms in ACT-R. In: Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Simulation., pp. 29–52. Cambridge University Press (2006)