Sometimes Less is More:

Why too many input languages can harm a system Peter Höfner

June 2019

DATA

61

www.csiro.au



Theorem: We developed an efficient algorithm to solve a complex problem using problem X, using the *Brisbane-methodology* and the *FMOz-formalism*.

Remark: The system *COOL* is implemented in the programming language PL, available at

http://fmoz_problemX.com.au/2019/



Proof: By simple induction





Remark: Our tool also supports the following input languages

- process algebras CCS, CSP, ABC
- Petri Nets
- Timed Automata
- Markov Decision Diagrams and Markow Chains

• ...

Therefore it is really user-friendly, isn't it?



Questions



Question Time



You accept all these input languages and transform them (internally) into the formalism FMOz, what guarantees can you give?







Further Discussion



- What structure does your translation have
 - branching bisimulation
 - strong/weak bisimulation
 - justness-preserving

- What properties can be checked
 - deadlock
 - livelock
 - timelock
 - safety properties
 - liveness properties
 - security/infoflow properties

Standard Reply





A Trivial Example



• From automata to some formal language



Is this a good translation?
 (CTL) Property:
 E G a

A Trivial Example



• From automata to some formal language



- Is this a good translation?
 (CTL) Property:
 A G a
- It is a (strong) simulation

A (Non-)Trivial Example



• From automata to some formal language



- It is a strong bisimulation (roughly)
- Is this a good translation? good for safety, not necessary for liveness (see Rob's talk@FMOz'18)



- Sometimes Less is More
- "Do not just add further input languages to your tool"
- If you do, state which properties can be shown (and prove it)
- ask others!



Bisimulation: A Formal Definition



 standard technique to compare labelled transition systems (LTSs)



A binary relation $\mathcal{R} \subseteq S_1 \times S_2$ is a *strong simulation* between two transition systems if it satisfies, for all a

if $p \mathcal{R} q$ and $p \xrightarrow{a}_{1} p'$ then $\exists q'. q \xrightarrow{a}_{2} q'$ and $p' \mathcal{R} q'$.

A bisimulation is a relation \mathcal{R} with both \mathcal{R} and \mathcal{R}^{\vee} being strong simulations.

It's Obvious: A Case Study



- From AWN (Algebra for Wireless Networks) to mCRL2 (milli Common Representation Language)
- why: AWN is ideal for reasoning about routing protocols (see FMOz'18) mCRL2 provides more than 50 tools (model checker, simulator, analyser, ...) http://mcrl2.org

AWN vs mCRL2



AWN (process algebra)	mCRL2 (process algebra)
choice (+), composition (.),	choice (+), composition (.),
data structure	data structure
process call	process calloth
creating a bisimard creating bisimard is abaightforward	

AWN vs mCRL2



AWN (process algebra)	mCRL2 (process algebra)
choice (+), composition (.),	choice (+), composition (.),
data structure	data structure
process call	process calloth
	ion betwee
layered (nodes, networks,) inula	single layer
4 different synchronisation from	1 operator
dynamic topologies	???

AWN to mCRL2: A Summary



- 1 Master's Thesis (of a very smart student)
- 12 months of work
- 1 conference publication
- 60+ pages proof
- required the introduction of 2 new variants of bisimulation
- we did several mistakes in our translation that were only discovered through formal proof
- But yes, a translation from any input language is obvious and straightforward



The formal translation (most parts)

 $T_V(\zeta, \mathbf{broadcast}(ms).P) = \sum_{\mathsf{D}:\mathsf{Set}(\mathsf{IP})} \mathbf{cast}(\mathsf{IP}, \mathsf{D}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{groupcast}(dests, ms).P) = \sum_{\mathsf{D}: Set(IP)} \mathbf{cast}(dests^{\zeta}, \mathsf{D}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{unicast}(dest, ms). P \triangleright Q) = \mathbf{cast}(\{dest^{\zeta}\}, \{dest^{\zeta}\}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $+ \neg \mathbf{uni}(\{dest^{\zeta}\}, \emptyset, ms^{\zeta}) \cdot T_V(\zeta, Q)$ $T_V(\zeta, \mathbf{send}(ms).P) = \mathbf{send}(\emptyset, \emptyset, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \operatorname{\mathbf{deliver}}(data).P) = \sum_{ip:IP} \operatorname{\mathbf{del}}(ip, data^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{receive}(\mathtt{m}).P) = \sum_{\mathtt{D}, \mathtt{D}': \mathtt{Set}(\mathtt{IP})} \mathbf{receive}(\mathtt{D}, \mathtt{D}', \mathtt{m}) \cdot T_{V \cup \{\mathtt{m}\}}(\zeta^{\setminus \mathtt{m}}, P)$ $T_V(\zeta, \llbracket v := exp \rrbracket P) = \sum_{v:sort(v)} (v = exp^{\zeta}) \rightarrow$ $(\sum_{\mathbf{v}: \text{sort}(\mathbf{v})} (\mathbf{v} = \mathbf{y}) \to \mathbf{t} \cdot \mathbf{T}_{V \cup \{\mathbf{v}\}} (\zeta^{\setminus \mathbf{v}}, P))$ $T_V(\zeta, X(exp_1, \cdots, exp_n)) = X(exp_1^{\zeta}, \cdots, exp_n^{\zeta})$ $T_V(\zeta, P+Q) = T_V(\zeta, P) + T_V(\zeta, Q)$ $T_V(\zeta, [\varphi]P) = \sum_{F_V(\varphi) \setminus V} \varphi^{\zeta} \to \mathbf{t} \cdot T_{V \cup F_V(\varphi)}(\zeta, P)$



the translation (the ugly parts)

 $T(ip:P:R) = \nabla_V \Gamma_C(T(P) \| G(ip,R))$ where $V = \{t, starcast, arrive, deliver, connect, disconnect\}$ where $C = \{ cast | cast \rightarrow starcast, \neg uni | \neg uni \rightarrow t, \}$ $del|\overline{del} \rightarrow deliver, receive|\overline{receive} \rightarrow arrive\}$ where $G(ip, R) \stackrel{\text{def}}{=} \sum_{D,D': \text{Set}(IP)} (R \cap D = D') \rightarrow \overline{cast}(D, D', m) \cdot G(ip, R)$ $+ \sum_{\substack{\mathsf{d:IP}\\\mathsf{m:MSG}}}^{\mathsf{m:MSG}} (\mathsf{d} \notin \mathsf{R}) \to \overline{\neg \mathbf{uni}}(\{\mathsf{d}\}, \emptyset, \mathsf{m}) \cdot G(\mathtt{ip}, \mathsf{R}) \\ + \sum_{\mathtt{data:DATA}} \overline{\mathbf{del}}(\mathtt{ip}, \mathtt{data}) \cdot G(\mathtt{ip}, \mathsf{R})$ + $\sum_{ip':IP} \operatorname{connect}(ip, ip') \cdot G(ip, \mathsf{R} \cup \{ip'\})$ + $\sum_{ip':IP} \mathbf{connect}(ip', ip) \cdot G(ip, \mathtt{R} \cup \{ip'\})$ $+ \sum_{\texttt{ip}',\texttt{ip}'':\text{IP}} (\texttt{ip} \notin \{\texttt{ip}',\texttt{ip}''\}) \rightarrow \textbf{connect}(\texttt{ip}',\texttt{ip}'') \cdot G(\texttt{ip},\texttt{R})$ + $\sum_{ip':IP} \operatorname{disconnect}(ip, ip') \cdot G(ip, \mathbb{R} \setminus \{ip'\})$ + $\sum_{ip':IP} \operatorname{disconnect}(ip', ip) \cdot G(ip, \mathbb{R} \setminus \{ip'\})$ $+ \sum_{\texttt{ip}',\texttt{ip}'':\text{IP}} (\texttt{ip} \notin \{\texttt{ip}',\texttt{ip}''\}) \rightarrow \textbf{disconnect}(\texttt{ip}',\texttt{ip}'') \cdot G(\texttt{ip},\texttt{R})$ $+ \sum_{{\tt D},{\tt D}':{\rm Set}({\rm IP})\atop {\tt m:MSG}} ({\tt ip}\in{\tt D}') \rightarrow \overline{{\tt receive}}({\tt D},{\tt D}',{\tt m}) \cdot G({\tt ip},{\tt R})$ $+ \sum_{\mathtt{D},\mathtt{D}':\mathtt{Set}(\mathtt{IP})}^{\mathtt{max}} (\mathtt{ip} \notin \mathtt{D}') \to \mathbf{arrive}(\mathtt{D},\mathtt{D}',\mathtt{m}) \cdot G(\mathtt{ip},\mathtt{R})$



The formal translation (most parts)

 $T_V(\zeta, \mathbf{broadcast}(ms).P) = \sum_{\mathsf{D}:\mathsf{Set}(\mathsf{IP})} \mathbf{cast}(\mathsf{IP}, \mathsf{D}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{groupcast}(dests, ms).P) = \sum_{\mathsf{D}: Set(IP)} \mathbf{cast}(dests^{\zeta}, \mathsf{D}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{unicast}(dest, ms). P \triangleright Q) = \mathbf{cast}(\{dest^{\zeta}\}, \{dest^{\zeta}\}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $+ \neg \mathbf{uni}(\{dest^{\zeta}\}, \emptyset, ms^{\zeta}) \cdot T_V(\zeta, Q)$ $T_V(\zeta, \mathbf{send}(ms).P) = \mathbf{send}(\emptyset, \emptyset, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{deliver}(data).P) = \sum_{ip:IP} \mathbf{del}(ip, data^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{receive}(\mathtt{m}).P) = \sum_{\mathtt{D}, \mathtt{D}': \mathtt{Set}(\mathtt{IP})} \mathbf{receive}(\mathtt{D}, \mathtt{D}', \mathtt{m}) \cdot T_{V \cup \{\mathtt{m}\}}(\zeta^{\setminus \mathtt{m}}, P)$ $T_V(\zeta, \llbracket v := exp \rrbracket P) = \sum_{v:sort(v)} (y = exp^{\zeta}) \rightarrow$ $(\sum_{\mathbf{v}:\operatorname{sort}(\mathbf{v})}(\mathbf{v}=\mathbf{y}) \to \mathbf{t} \cdot \mathrm{T}_{V \cup \{\mathbf{v}\}}(\zeta^{\setminus \mathbf{v}}, P))$ $T_V(\zeta, X(exp_1, \cdots, exp_n)) = X(exp_1^{\zeta}, \cdots, exp_n^{\zeta})$ $T_V(\zeta, P+Q) = T_V(\zeta, P) + T_V(\zeta, Q)$ $T_V(\zeta, [\varphi]P) = \sum_{FV(\varphi) \setminus V} \varphi^{\zeta} \to \mathbf{t} \cdot T_{V \cup FV(\varphi)}(\zeta, P)$

Bisimulation modulo Renaming





A binary relation $\mathcal{R} \subseteq S_1 \times S_2$ is a strong simulation modulo renaming between two transition systems for a bijective renaming function f, if, for all a

if
$$p \mathcal{R} q$$
 and $p \xrightarrow{a}_{1} p'$ then $\exists q'. q \xrightarrow{f(a)}_{2} q'$ and $p' \mathcal{R} q'$.

A bisimulation modulo renaming is a relation \mathcal{R} with both \mathcal{R} and \mathcal{R}^{\vee} being strong simulations modulo renaming.

Bisimulation and Data Congruence



- we also had to take data congruence into account (actions a(3), a(1+2) and a(6-3) should be treated the way)
- luckily, data congruences are "standard"
- in our setting data concurrence (\equiv) is a strong bisimulation
- strong bisimulations are compositional



The formal translation (most parts)

 $T_V(\zeta, \mathbf{broadcast}(ms).P) = \sum_{\mathsf{D}:\mathsf{Set}(\mathsf{IP})} \mathbf{cast}(\mathsf{IP}, \mathsf{D}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{groupcast}(dests, ms).P) = \sum_{\mathsf{D}: Set(IP)} \mathbf{cast}(dests^{\zeta}, \mathsf{D}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{unicast}(dest, ms)) \ge Q \ge \mathbf{cast}(\{dest^{\zeta}\}, \{dest^{\zeta}\}, ms^{\zeta}) \cdot T_V(\zeta, P)$ $+ \neg \mathbf{uni}(\{dest^{\varsigma}\}, \emptyset, ms^{\varsigma}) \cdot \mathbf{T}_{V}(\zeta, Q)$ $T_V(\zeta, \mathbf{send}(ms).P) = \mathbf{send}(\emptyset, \emptyset, ms^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{deliver}(data).P) = \sum_{ip:IP} \mathbf{del}(ip, data^{\zeta}) \cdot T_V(\zeta, P)$ $T_V(\zeta, \mathbf{receive}(\mathtt{m}).P) = \sum_{\mathtt{D}, \mathtt{D}': \mathtt{Set}(\mathtt{IP})} \mathbf{receive}(\mathtt{D}, \mathtt{D}', \mathtt{m}) \cdot T_{V \cup \{\mathtt{m}\}}(\zeta^{\setminus \mathtt{m}}, P)$ $T_V(\zeta, \llbracket v := exp \rrbracket P) = \sum_{v:sort(v)} (y = exp^{\zeta}) \rightarrow$ $(\sum_{\mathbf{v}:sort(\mathbf{v})} (\mathbf{v} = \mathbf{y}) \to \mathbf{t} \cdot T_{V \cup \{\mathbf{v}\}}(\zeta^{\setminus \mathbf{v}}, P))$ $T_V(\zeta, X(exp_1, \cdots, exp_n)) = X(exp_1^{\zeta}, \cdots, exp_n^{\zeta})$ $T_V(\zeta, P+Q) = T_V(\zeta, P) + T_V(\zeta, Q)$ $T_V(\zeta, [\varphi]P) = \sum_{FV(\varphi) \setminus V} \varphi^{\zeta} \to \mathbf{t} \cdot T_{V \cup FV(\varphi)}(\zeta, P)$

Warped Bisimulation



 $\exists q', (\{a\}, \{b_1, b_2, b_3, b_4\}) \in \mathcal{A}.$

'R

 the layered structure of AWN required another generalisation



A binary relation $\mathcal{R} \subseteq S_1 \times S_2$ is a \mathcal{A} -warped simulation up to \equiv between two transition systems for a relation \mathcal{A} (between sets of action labels) if it satisfies

if
$$p \mathcal{R} q$$
 and $p \xrightarrow{a}_1 p''$ then $\exists A_1, A_2, p', q'$.
 $(a \in A_1, p'' \equiv p', A_1 \mathcal{A} A_2, p \xrightarrow{A_1}_1 \equiv p', q \xrightarrow{A_2}_2 \equiv q' \text{ and } p' \mathcal{R} q')$,

where $p \xrightarrow{\mathsf{A}}_1 \equiv p' : \Leftrightarrow \forall a \in \mathsf{A}. \exists p''. p \xrightarrow{a}_1 p'' \land p'' \equiv p'.$

An \mathcal{A} -warped bisimulation up to \equiv is a relation \mathcal{R} with \mathcal{R} is an \mathcal{A} -warped bisimulation up to \equiv and $\mathcal{R} \subset \mathcal{A} \subset$ -warped bisimulation up to \equiv .

Conclusion



- we have brilliant tools that give formal guarantees
- adding another input language carelessly may lead
 - false results
 - reduces confidence in formal methods and formal tools
 - reduces our credibility

Future Work



- keep asking (raise awareness for the problem)
- how to scale the proof effort (our case study were 2 process algebras, no time, no probabilities)
- what structure is needed to maintain properties
 - safety = (strong) bisimulation
 - liveness = justness (or fairness) ?
 - security = ???
- is there a generic way to prove these translations (I doubt)

Thank you

DATA

61

Data61 Peter Höfner

- t +61 2 9490 5861
- e peter.hoefner@data61.csiro.au
- w www.data61.csiro.au



www.csiro.au