

## FAIR SIMULATION FOR NONDETERMINISTIC AND PROBABILISTIC BÜCHI AUTOMATA: A COALGEBRAIC PERSPECTIVE

NATSUKI URABE\* AND ICHIRO HASUO

Dept. Computer Science, The University of Tokyo, Hongo 7-3-1, Tokyo 113-8656, Japan  
*e-mail address:* urabenatsuki@is.s.u-tokyo.ac.jp

National Institute of Informatics, Hitotsubashi 2-1-2, Tokyo 101-8430, Japan  
*e-mail address:* i.hasuo@acm.org

*In honor of Jiří Adámek on the occasion of his seventieth birthday*

**ABSTRACT.** Notions of *simulation*, among other uses, provide a computationally tractable and sound (but not necessarily complete) proof method for language inclusion. They have been comprehensively studied by Lynch and Vaandrager for nondeterministic and timed systems; for *Büchi* automata the notion of *fair simulation* has been introduced by Henzinger, Kupferman and Rajamani. We contribute to a generalization of fair simulation in two different directions: one for nondeterministic *tree* automata previously studied by Bomhard; and the other for *probabilistic* word automata with finite state spaces, both under the Büchi acceptance condition. The former nondeterministic definition is formulated in terms of systems of fixed-point equations, hence is readily translated to parity games and is then amenable to Jurdziński’s algorithm; the latter probabilistic definition bears a strong ranking-function flavor. These two different-looking definitions are derived from one source, namely our *coalgebraic* modeling of Büchi automata. Based on these coalgebraic observations, we also prove their soundness: a simulation indeed witnesses language inclusion.

### 1. INTRODUCTION

Notions of *simulation*—typically defined as a binary relation subject to a coinductive “one-step mimicking” condition—have been studied extensively in formal verification and process theory. Sometimes existence of a simulation itself is interesting—taking it as the definition of an abstraction/refinement relationship—but another notable use is as a *proof method* for language inclusion. Language inclusion is fundamental in model checking but often hard to check itself; looking for a simulation—which *witnesses* language inclusion, by its *soundness* property, in a step-wise manner—is then a sound (but generally not complete) alternative. For example, (finite) language inclusion between weighted automata

2012 ACM CCS: [Theory of computation]: Logic—Verification by model checking.

*Key words and phrases:* Büchi automaton, fair simulation, tree automaton, probabilistic automaton, coalgebra .

\* JSPS Research Fellow.

with weights in the semiring of real numbers with ordinary addition and multiplication is undecidable [BC03], while existence of certain simulations is PTIME, see [UH17].

Simulation notions have been introduced for many different types of systems: nondeterministic [LV95], timed [LV96] and probabilistic [JL91], among others. Conventionally many studies take the trivial acceptance condition (any run that does not diverge, i.e. that does not come to a deadend, is accepted). Recently, however, there have been several works on simulations under the *Büchi* and *parity* acceptance conditions [HKR02, EWS05, FW06]. In such settings a simulation notion is subject to an (inevitable) nonlocal *fairness* condition (on top of the local condition of one-step mimicking); and often a fair simulation is characterized as a winning strategy of a suitable *parity game*, which is then searched using Jurdziński's algorithm [Jur00].

Like simulation notions for the trivial acceptance condition, fair simulation can be used for proving language inclusion. Moreover, it also has a logical characterization: there exists a certain universal fragment of the alternation-free  $\mu$ -calculus such that for two systems  $\mathcal{X}$  and  $\mathcal{Y}$ , there exists a fair simulation from  $\mathcal{X}$  to  $\mathcal{Y}$  if and only if all the formulas in the fragment that are satisfied in  $\mathcal{Y}$  are also satisfied in  $\mathcal{X}$  [HKR02]. In contrast, differently from many other simulation notions, fair simulations cannot be used for state-space reduction [EWS05]. In [EWS05], a weaker simulation notion called *delayed simulation* is introduced for this purpose.

**1.1. Contributions.** It is in the context of fair simulation for word/tree automata with nondeterministic/probabilistic branching that the current paper contributes:

- (1) We define fair simulation for *nondeterministic tree* automata with the Büchi acceptance condition. We express the notion using a *system of fixed-point equations*—with explicit  $\mu$ 's and  $\nu$ 's indicating least or greatest—and thus the definition makes sense for infinite-state automata too. We also interpret it in terms of a parity game, which is subject to an algorithmic search when the problem instance is finitary. The resulting parity game essentially coincides with the one in [vB08].
- (2) We define fair simulation for *probabilistic word* automata with the Büchi acceptance condition, this time with the additional condition that on the simulating side we have a finite-state automaton. This simulation notion is given by a *matrix* (instead of a relation); this follows our previous work [UH17] that uses *linear programming* to search for such a matrix simulation. Our current notion also requires suitable *approximation sequences* for witnessing well-foundedness, with a similar intuition to *ranking functions*.

For the former *nondeterministic tree* setting ((1) in the above), a notion of fair simulation—in addition to direct and delayed simulation—has already been introduced in [vB08]. Their definition focuses on finite state spaces and is formulated using a parity game. In contrast, our notion is given in terms of fixed-point equations—it generalizes the  $\mu$ -calculus characterization of fair simulations from words (see e.g. [JP06]) to trees. An obvious advantage of our fixed-point characterization over the parity game-based one is that ours makes sense for infinite state systems too. For the latter *probabilistic word* setting ((2) in the above), we introduce fair simulation for the first time (to the best of our knowledge).

In both settings our main technical result is *soundness*, that is, existence of a fair simulation implies trace inclusion. We also exhibit nontrivial examples of fair simulations.

**1.2. Theoretical Backgrounds.** Our two simulation notions (for nondeterministic tree automata and probabilistic word ones) look rather different, but they are derived from the same theoretical insights. The insights come from: 1) the theory of *coalgebra* [Jac16, Rut00], in particular the generic *Kleisli theory of trace and simulation* [Jac04, Cîr10, Has06, UH17, UH15]; and 2) our recent work [HSC16] on a lattice-theoretic foundation of *nested/alternating fixed points*, where we generalize *progress measures*, a central notion in Jurdiński’s algorithm [Jur00] for parity games. We rely on both of these series of work also for soundness proofs, where we follow yet another recent work of ours [USH16] in which we characterize the accepted language of a Büchi automaton by an “equational system” of diagrams in a Kleisli category. In this paper we shall briefly describe these general theories behind our current results, focusing on their instances that are relevant.

**1.3. A Tribute to Jiří Adámek.** Working in coalgebraic modeling of dynamical systems and their analysis, we owe to Jiří almost the field itself on which we stand.

His earlier works like [Adá74, AK79] paved the way to one of the most fundamental ideas in the field, namely the identification of final/terminal coalgebras as categorical fully abstract domains of behaviors of non-terminating systems. Fixed-point equations—to which a final coalgebra is the greatest categorical solution in case an equation is given by a functor—have been a central theme in his more recent works too. These include the line of work pursued in [AMV11] and many others, where “solution operators” for fixed-point equations and their axioms are studied in an elegantly categorical fashion. All these works of his have been a great source of inspiration for us.

We wish to dedicate the current work to Jiří. It continues our recent line of work [HSC16, USH16, UHH17] in which we pursue: categorical understanding of nested and alternating least and greatest fixed-point equations, and proof methods for such fixed-point specifications. Categorical solutions to fixed-point equations play a central role here, much like in Jiří’s series of work, while we believe our use of orders between arrows and explicit  $\mu$ ’s and  $\nu$ ’s (like in (5.3)) is a crucial step ahead towards accommodating complex fixed-point specifications (like persistence and recurrence) in the categorical study of coalgebras. More specifically we contribute fair simulation notions as witnesses for Büchi language inclusion. Hopefully our results demonstrate potential practical values of mathematical and categorical understanding of systems, a theme Jiří has been pursuing throughout his career.

**1.4. Organization of the Paper.** In Section 2 we introduce *equational systems*, essentially fixing notations for alternating greatest and least fixed points. These notations—and the idea that fixed-point equations play important roles—are used in Sections 3–4 where we concretely describe: our system models; their accepted languages; simulation definitions; and soundness results. In this paper we consider *nondeterministic tree* automata and *probabilistic word* automata, both with the Büchi acceptance condition, as system models. Up to this point everything is in set-theoretic terms, without category theory.

The rest of the paper is devoted to soundness proofs and the theoretical perspectives behind. In Section 5 we review the coalgebraic backgrounds: Kleisli categories, coalgebras, trace semantics [Jac04, HJS07, Cîr10], simulations (under the trivial acceptance condition) [Has06, UH15], and coalgebraic trace for Büchi automata [USH16]. Finally, in Section 6 we take a coalgebraic look at simulations under the Büchi condition: our first attempt (fair simulation *with dividing*) is sound but not practically desirable; we show how

we can circumvent this additional construct of dividing, and how we can obtain the concrete definitions in Sections 3–4.

In Section 7 we conclude and suggest some directions of future work.

## 2. PRELIMINARIES: EQUATIONAL SYSTEMS

*Nested, alternating* greatest and least fixed points—as in a  $\mu$ -calculus formula  $\nu u_2.\mu u_1.(p \wedge u_2) \vee \Box u_1$ —are omnipresent in specification and verification. For their relevance to the Büchi acceptance condition one can recall the well-known translation of LTL formulas to Büchi automata and vice versa (see [Var95] for example). To express such fixed points we follow [CKS92, AN01] and use *equational systems*—instead of textual  $\mu$ -calculus-like presentations.

**Definition 2.1** (equational system). Let  $L_1, \dots, L_m$  be posets. We write  $\sqsubseteq$  for the orders over the posets. An *equational system*  $E$  over  $L_1, \dots, L_m$  is an expression

$$u_1 =_{\eta_1} f_1(u_1, \dots, u_m), \quad \dots, \quad u_m =_{\eta_m} f_m(u_1, \dots, u_m) \quad (2.1)$$

where:  $u_1, \dots, u_m$  are *variables*,  $\eta_1, \dots, \eta_m \in \{\mu, \nu\}$ , and  $f_i: L_1 \times \dots \times L_m \rightarrow L_i$  is a monotone function. A variable  $u_j$  is a  $\mu$ -*variable* if  $\eta_j = \mu$ ; it is a  $\nu$ -*variable* if  $\eta_j = \nu$ .

**2.1. Solutions of Equational Systems.** In this section we define the *solution* of an equational system. For the equational system  $E$  in Def. 2.1, its solution is defined as a family  $(l_1^{\text{sol}}, \dots, l_m^{\text{sol}}) \in L_1 \times \dots \times L_m$ . We first briefly sketch its definition.

We assume that  $L_i$ 's have enough suprema and infima. The definition proceeds as follows: 1) we solve the first equation of (2.1) for  $u_1$  to obtain an interim solution  $u_1 = l_1^{(1)}(u_2, \dots, u_m)$  that is parameterized by  $u_2, \dots, u_m$ ; 2) it is used in the second equation to eliminate  $u_1$  and yield a new equation  $u_2 =_{\eta_2} f_2^\ddagger(u_2, \dots, u_m)$ ; 3) solving it again gives an interim solution  $u_2 = l_2^{(2)}(u_3, \dots, u_m)$ ; 4) continuing this way from left to right eventually eliminates all the variables and leads to a closed solution  $u_m = l_m^{(m)} \in L_m$ ; and 5) by propagating these closed solutions back from right to left, we obtain closed solutions for all of  $u_1, \dots, u_m$ . To summarize, when we are solving the  $i$ -th equation  $u_i =_{\eta_i} f_i(u_1, \dots, u_m)$ , we first substitute  $u_1, \dots, u_{i-1}$  with the current interim solutions  $l_1^{(i-1)}(u_i, \dots, u_m), \dots, l_{i-1}^{(i-1)}(u_i, \dots, u_m)$ , and solve the equation for  $u_i$ , regarding  $u_{i+1}, \dots, u_m$  as parameters. We give now a formal definition.

**Definition 2.2** (solution). Let  $E$  be the equational system in Definition 2.1. For each  $i \in [1, m]$  and  $j \in [1, i]$ , we define monotone functions  $f_i^\ddagger: L_i \times \dots \times L_m \rightarrow L_i$  and  $l_j^{(i)}: L_{i+1} \times \dots \times L_m \rightarrow L_j$  by induction on  $i$  as follows.

- When  $i = 1$ ,

$$f_1^\ddagger(l_1, \dots, l_m) := f_1(l_1, \dots, l_m), \quad \text{and}$$

$$l_1^{(1)}(l_2, \dots, l_m) := \begin{cases} l_{\text{lfp}} & (\eta_1 = \mu \text{ and } f_1^\ddagger(\_, l_2, \dots, l_m) \text{ has the lfp } l_{\text{lfp}} \in L_1) \\ l_{\text{gfp}} & (\eta_1 = \nu \text{ and } f_1^\ddagger(\_, l_2, \dots, l_m) \text{ has the gfp } l_{\text{gfp}} \in L_1) \\ \text{undefined} & (\text{otherwise}). \end{cases}$$

Note here that completeness of  $L_1$  is not assumed and therefore the monotone function  $f_1^\ddagger(\_, l_2, \dots, l_m): L_1 \rightarrow L_1$  does not necessarily have the lfp or gfp.

- For the step case, the function  $f_{i+1}^\ddagger$  is defined using the  $i$ -th interim solutions  $l_1^{(i)}, \dots, l_i^{(i)}$  for the variables  $u_1, \dots, u_i$  obtained so far:

$$f_{i+1}^\ddagger(l_{i+1}, \dots, l_m) := \begin{cases} f_{i+1}(l_1^{(i)}(l_{i+1}, \dots, l_m), \dots, l_i^{(i)}(l_{i+1}, \dots, l_m), l_{i+1}, \dots, l_m) \\ \quad (l_j^{(i)}(l_{i+1}, \dots, l_m) \text{ is defined for each } j \in [1, i]) \\ \text{undefined} \quad \text{(otherwise)}. \end{cases}$$

For  $j = i + 1$ ,  $l_j^{(i+1)}$  is defined by

$$l_{i+1}^{(i+1)}(l_{i+2}, \dots, l_m) := \begin{cases} l_{\text{lfp}} & \left( \eta_{i+1} = \mu, \text{ and } \right. \\ & \left. f_{i+1}^\ddagger(\_, l_{i+2}, \dots, l_m) \text{ has the lfp } l_{\text{lfp}} \in L_{i+1} \right) \\ l_{\text{gfp}} & \left( \eta_{i+1} = \nu, \text{ and } \right. \\ & \left. f_{i+1}^\ddagger(\_, l_{i+2}, \dots, l_m) \text{ has the gfp } l_{\text{gfp}} \in L_{i+1} \right) \\ \text{undefined} & \text{(otherwise)}. \end{cases}$$

For  $j \in [1, i]$ ,  $l_j^{(i+1)}$  is defined using  $l_{i+1}^{(i+1)}(l_{i+2}, \dots, l_m)$  as follows.

$$l_j^{(i+1)}(l_{i+2}, \dots, l_m) := \begin{cases} l_j^{(i)}(l_{i+1}^{(i+1)}(l_{i+2}, \dots, l_m), l_{i+2}, \dots, l_m) \\ \quad (l_{i+1}^{(i+1)}(l_{i+2}, \dots, l_m) \text{ is defined}) \\ \text{undefined} \quad \text{(otherwise)} \end{cases}$$

A family  $(l_1^{\text{sol}}, \dots, l_m^{\text{sol}}) \in L_1 \times \dots \times L_m$  is called the *solution* of  $E$  if  $l_j^{(m)} : 1 \rightarrow L_j$  is defined and  $l_j^{\text{sol}} = l_j^{(m)}(*)$  (here  $*$  is the unique element in  $1$ ) for each  $j \in [1, m]$ .

Note that the order of equations *matters*. For  $(u =_\mu v, v =_\nu u)$  the solution is  $u = v = \top$  while for  $(v =_\nu u, u =_\mu v)$  the solution is  $u = v = \perp$ . It is easy to see that all the functions  $f_i^\ddagger$  and  $l_j^{(i)}$  involved here are monotone. By the definition above, a solution exists if the function

$$f_i^\ddagger(\_, l_{i+1}, \dots, l_m) : L_i \longrightarrow L_i \quad (2.2)$$

in Definition 2.2 has both the least and the greatest fixed points for each  $i \in [1, m]$  and  $l_{i+1} \in L_{i+1}, \dots, l_m \in L_m$ . Their existence depends on how “complete” each  $L_i$  is and how “continuous” each  $f_i$  is. In the following proposition we present two sufficient conditions for existence of the least and the greatest fixed points.

**Proposition 2.3.** *Let  $E$  be the equational system in Definition 2.1. If either of the following conditions is satisfied, then  $E$  has a (necessarily unique) solution.*

- For each  $i \in [1, m]$ , the poset  $L_i$  is a complete lattice.
- For each  $i \in [1, m]$  we have the following.
  - $L_i$  has both the least and greatest elements.
  - $L_i$  is both  $\omega$ -complete and  $\omega^{\text{op}}$ -complete, that is, every increasing (or decreasing)  $\omega$ -chain has a supremum (or an infimum, respectively).
  - For each  $l_{i+1} \in L_{i+1}, \dots, l_m \in L_m$ , the function

$$f_i^\ddagger(\_, l_{i+1}, \dots, l_m) : L_i \longrightarrow L_i$$

in Definition 2.2 is both  $\omega$ -continuous and  $\omega^{\text{op}}$ -continuous, that is, the aforementioned suprema and infima are preserved by the function.  $\square$

If Condition (a) above is satisfied then existence of the least and the greatest fixed points of the function in (2.2) is ensured by the Knaster–Tarski theorem. In contrast, if Condition (b) is satisfied then existence of the least and the greatest fixed points is ensured by the Kleene fixed-point theorem.

As we will see later, Condition (a) is suitable for the nondeterministic setting while Condition (b) is suitable for the probabilistic setting.

**2.2. Progress Measure.** The notion of (lattice-theoretic) *progress measure* [HSC16], although not explicit, plays an important role in the current paper. We first briefly review its idea.

Verification of a fixed-point specification amounts mathematically to *underapproximating* the fixed point.<sup>1</sup> This is usually done very differently for gfp’s and lfp’s. For a gfp  $\nu f$  one provides an *invariant*  $l$ —a post-fixed point  $l \sqsubseteq f(l)$ —and then the Knaster-Tarski theorem yields  $l \sqsubseteq \nu f$ . However, for an lfp  $\mu f$ , the same argument (namely finding a pre-fixed point  $f(l) \sqsubseteq l$ ) would give an *overapproximation*; instead we should appeal to the Cousot-Cousot theorem [CC79] and consider the *approximation sequence*  $\perp \sqsubseteq f(\perp) \sqsubseteq \dots$ . The sequence eventually converges to  $\mu f$  (possibly after transfinite induction);<sup>2</sup> hence for every ordinal  $\alpha$ , the approximant  $f^\alpha(\perp)$  is an underapproximation of  $\mu f$ . This is the underlying principle of proofs by *ranking functions* of termination, for example.

Progress measures in [HSC16], generalizing the combinatorial notion of the same name in Jurziński’s algorithm for parity games [Jur00], are roughly combination of invariants and ranking functions. The latter two must be combined in an intricate manner so that they respect the order of equations in (2.1) (that is, priorities in parity games or  $\mu$ -calculus formulas); we do so with the help of a suitable truncated order.

Use of parity games is nowadays omnipresent, and the study of fair simulations is not an exception [EWS05]. Following those previous works, the basic idea behind our developments (below) is to generalize: parity games to equational systems (Definition 2.1); and accordingly, Jurziński’s (combinatorial) progress measure to our lattice-theoretic one [HSC16].

In the rest of this section we formally state the formal definition of progress measure, as well as its soundness and completeness results (against the solution of an equational system). To this end, we first review the notion of *prioritized ordinal*, which embodies the idea of *priority* in parity games. See [HSC16] for the relationship between the notion of prioritized ordinal and the notion of priority in parity games.

**Definition 2.4** (prioritized ordinal,  $\leq_i$ ). Let  $E$  be the equational system in (2.1) of Definition 2.1. Let us collect the indices of  $\mu$ -variables:  $\{i_1, \dots, i_k\} = \{i \in [1, m] \mid \eta_i = \mu \text{ in (2.1)}\}$ , and assume that  $i_1 < \dots < i_k$ . A *prioritized ordinal* for  $E$  is a  $k$ -tuple  $(\alpha_1, \dots, \alpha_k)$  of ordinals.

For each  $i \in [1, m]$  we define a preorder  $\leq_i$  between prioritized ordinals—called the  *$i$ -th truncated pointwise order*—as follows. If  $i_k < i$ , then  $(\alpha_1, \dots, \alpha_k) \leq_i (\alpha'_1, \dots, \alpha'_k)$  is always true. Otherwise, let  $a \in [1, k]$  be such that  $i_1 < \dots < i_{a-1} < i \leq i_a < \dots < i_k$ , that is,  $u_{i_a}$  is the  $\mu$ -variable with the smallest priority above that of  $i$ . Then we define  $(\alpha_1, \dots, \alpha_k) \leq_i (\alpha'_1, \dots, \alpha'_k)$  if we have  $\alpha_i \leq \alpha'_i$  for each  $i \in [a, k]$ .

<sup>1</sup>In some cases we might be interested in approximations with respect to *distance* rather than order [vBW05]. In such cases we can use the Banach fixed-point theorem instead of the Knaster-Tarski or Cousot-Cousot one.

<sup>2</sup>In case  $f$  is continuous the sequence converges after  $\omega$  steps. This is the Kleene fixed-point theorem.

**Definition 2.5** (progress measure for an equational system). Let  $E$  be the equational system in Definition 2.1. We further assume that for each  $i \in [1, m]$ ,  $L_i$  has the smallest element  $\perp$ . A *progress measure*  $p$  for  $E$  is given by a tuple  $p = ((\overline{\alpha}_1, \dots, \overline{\alpha}_k), (p_i(\alpha_1, \dots, \alpha_k))_{i, \alpha_1, \dots, \alpha_k})$  that consists of:

- the *maximum prioritized ordinal*  $(\overline{\alpha}_1, \dots, \overline{\alpha}_k)$ ; and
- the *approximants*  $p_i(\alpha_1, \dots, \alpha_k) \in L_i$ , defined for each  $i \in [1, m]$  and each prioritized ordinal  $(\alpha_1, \dots, \alpha_k)$  such that  $\alpha_1 \leq \overline{\alpha}_1, \dots, \alpha_k \leq \overline{\alpha}_k$ .

The approximants  $p_i(\alpha_1, \dots, \alpha_k)$  are subject to:

- (1) **(Monotonicity)** For each  $i \in [1, m]$ ,  $(\alpha_1, \dots, \alpha_k) \leq_i (\alpha'_1, \dots, \alpha'_k)$  implies  $p_i(\alpha_1, \dots, \alpha_k) \sqsubseteq p_i(\alpha'_1, \dots, \alpha'_k)$ .
- (2) **( $\mu$ -variables, base case)** Let  $a \in [1, k]$ . Then  $\alpha_a = 0$  implies  $p_{i_a}(\alpha_1, \dots, \alpha_a, \dots, \alpha_k) = \perp$ .
- (3) **( $\mu$ -variables, step case)** Let  $a \in [1, k]$ . Then there exist ordinals  $\beta_1, \dots, \beta_{a-1}$  such that  $\beta_1 \leq \overline{\alpha}_1, \dots, \beta_{a-1} \leq \overline{\alpha}_{a-1}$  and

$$p_{i_a}(\alpha_1, \dots, \alpha_{a-1}, \alpha_a + 1, \alpha_{a+1}, \dots, \alpha_k) \sqsubseteq f_{i_a} \left( \begin{array}{c} p_1(\beta_1, \dots, \beta_{a-1}, \alpha_a, \alpha_{a+1}, \dots, \alpha_k), \\ \dots, \\ p_m(\beta_1, \dots, \beta_{a-1}, \alpha_a, \alpha_{a+1}, \dots, \alpha_k) \end{array} \right). \quad (2.3)$$

- (4) **( $\mu$ -variables, limit case)** Let  $a \in [1, k]$  and let  $\alpha_a$  be a limit ordinal. Then the supremum  $\bigsqcup_{\beta < \alpha_a} p_{i_a}(\alpha_1, \dots, \beta, \dots, \alpha_k) \in L_{i_a}$  exists and we have:

$$p_{i_a}(\alpha_1, \dots, \alpha_a, \dots, \alpha_k) \sqsubseteq \bigsqcup_{\beta < \alpha_a} p_{i_a}(\alpha_1, \dots, \beta, \dots, \alpha_k). \quad (2.4)$$

- (5) **( $\nu$ -variables)** Let  $i \in [1, m] \setminus \{i_1, \dots, i_k\}$ ; and let  $a \in [1, k]$  be such that  $i_1 < \dots < i_{a-1} < i < i_a < \dots < i_k$ . Let  $(\alpha_1, \dots, \alpha_k)$  be a prioritized ordinal. Then there exist ordinals  $\beta_1, \dots, \beta_{a-1}$  such that  $\beta_1 \leq \overline{\alpha}_1, \dots, \beta_{a-1} \leq \overline{\alpha}_{a-1}$  and

$$p_i(\alpha_1, \dots, \alpha_{a-1}, \alpha_a, \dots, \alpha_k) \sqsubseteq f_i \left( \begin{array}{c} p_1(\beta_1, \dots, \beta_{a-1}, \alpha_a, \dots, \alpha_k), \\ \dots, \\ p_m(\beta_1, \dots, \beta_{a-1}, \alpha_a, \dots, \alpha_k) \end{array} \right). \quad (2.5)$$

The definition combines the features of ranking functions (Conditions 2–4) and those of invariants (Condition 5). Note also that in each clause ordinals with smaller priorities can be modified to arbitrary  $\beta_i$ .

**Remark 2.6.** The definition of a progress measure in Definition 2.5 is slightly different from the one in [HSC16], in the following points.

- (1) Condition (1) is given using the truncated *pointwise* order instead of the truncated *lexicographic* order.
- (2) It is not assumed that each  $L_i$  is a complete lattice. Instead, in Condition (4), existence of the supremum is explicitly required.

The difference (1) is made for the sake of cleanliness of the soundness proof for our notion of simulation (Theorem 6.13). The difference (2) is made because, in the probabilistic setting (see e.g. Example 6.5), we should consider progress measures where each  $L_i$  is not a complete lattice or even a depo. Because of the latter difference, in the correctness theorem below, we need extra assumptions ((i) and (ii)) that do not appear in the correctness theorem in [HSC16].

Despite these modifications, the notion of progress measure in Definition 2.5 shares correctness properties with the original definition in [HSC16]—soundness and completeness. The proofs are almost the same as the ones in [HSC16].

**Theorem 2.7** (correctness of progress measures). *Let  $E$  be the equational system (2.1) and assume that  $E$  has the solution  $(l_1^{sol}, \dots, l_m^{sol})$ . We further assume that for each  $i \in [1, m]$ ,*

- (i) *the poset  $L_i$  has the least element  $\perp$  and is  $\omega$ -complete; and*
- (ii) *for each  $l_{i+1} \in L_{i+1}, \dots, l_m \in L_m$ , the function*

$$f_i^\ddagger(\_, l_{i+1}, \dots, l_m) : L_i \longrightarrow L_i$$

*in Definition 2.2 is  $\omega$ -continuous.*

Then we have the following.

- (1) (**Soundness**) *For each progress measure  $p = ((\overline{\alpha}_1, \dots, \overline{\alpha}_k), (p_i(\alpha_1, \dots, \alpha_k))_{i, \alpha_1, \dots, \alpha_k})$  we have  $p_i(\overline{\alpha}_1, \dots, \overline{\alpha}_k) \sqsubseteq l_i^{sol}$  for each  $i \in [1, m]$ .*
- (2) (**Completeness**) *There exists a progress measure  $p = ((\overline{\alpha}_1, \dots, \overline{\alpha}_k), (p_i(\alpha_1, \dots, \alpha_k))_{i, \alpha_1, \dots, \alpha_k})$  that achieves the solution, that is,  $p_i(\overline{\alpha}_1, \dots, \overline{\alpha}_k) = l_i^{sol}$  for each  $i \in [1, m]$ . Moreover we can find  $p$  such that  $\overline{\alpha}_i \leq \omega$  for each  $i \in [1, m]$ .  $\square$*

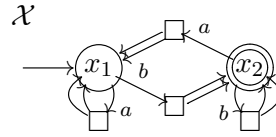
### 3. FAIR SIMULATION FOR NONDETERMINISTIC BÜCHI TREE AUTOMATA

A *ranked alphabet* is a set  $\Sigma$  with a function  $|\_ | : \Sigma \rightarrow \mathbb{N}$  that gives an *arity* to each  $\sigma \in \Sigma$ .

**Definition 3.1** (NBTA). A *nondeterministic Büchi tree automaton* (NBTA) is given by a quintuple  $\mathcal{X} = (X, \Sigma, \delta, I, \text{Acc})$  consisting of a state space  $X$ , a ranked alphabet  $\Sigma$ , a *transition function*  $\delta : X \rightarrow \mathcal{P}(\prod_{\sigma \in \Sigma} X^{|\sigma|})$ , a set  $I \subseteq X$  of the *initial states*, and a set  $\text{Acc} \subseteq X$  of the *accepting states* (often designated by  $\odot$ ).

**Example 3.2.** We define an NBTA  $\mathcal{X} = (X, \Sigma, \delta, I, \text{Acc})$  as follows.

- $X = \{x_1, x_2\}$
- $\Sigma = \{a, b\}$  where  $|a| = |b| = 2$
- $\delta(x_1) = \delta(x_2) = \{(a, (x_1, x_1)), (b, (x_2, x_2))\}$
- $I = \{x_1\}$
- $\text{Acc} = \{x_2\}$



Then  $\mathcal{X}$  can be illustrated as in the above. Here  $x \xrightarrow{\sigma} \square \rightrightarrows_z^y$  denotes  $(\sigma, (y, z)) \in \delta(x)$ .

**3.1. Accepted Languages of Nondeterministic Büchi Tree Automata.** We start with reviewing necessary notions for defining accepted (tree) languages of NBTAs. They are all as usual.

**Notation 3.3.** We let  $\mathbb{N}^*$  and  $\mathbb{N}^\omega$  denote the sets of finite and infinite sequences over natural numbers, respectively. Moreover we let  $\mathbb{N}^\infty := \mathbb{N}^* \cup \mathbb{N}^\omega$ . Concatenation of finite/infinite sequences, and/or characters are denoted simply by juxtaposition. Given an infinite sequence  $\pi = \pi_1 \pi_2 \dots \in \mathbb{N}^\omega$  (here  $\pi_i \in \mathbb{N}$ ), its prefix  $\pi_1 \dots \pi_n$  is denoted by  $\pi_{\leq n}$ .

The following formalization of trees and related notions are standard, with its variations used in [CHS14] for example.



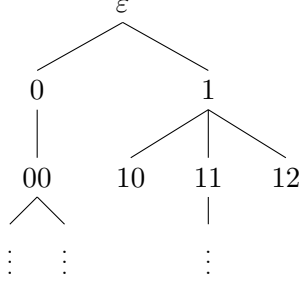
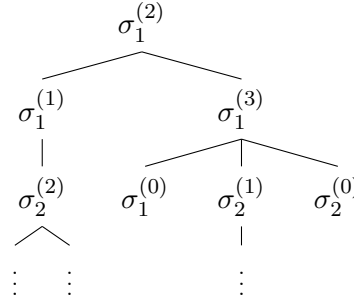


Figure 1: Positions in a tree


 Figure 2: The arity of a label, and the number of successors. Here  $\sigma_j^{(i)} \in \Sigma$  is assumed to be of arity  $i$ .

**Definition 3.4** ( $\Sigma$ -tree). Let  $\Sigma$  be a ranked alphabet, with each element  $\sigma \in \Sigma$  coming with its arity  $|\sigma| \in \mathbb{N}$ . A  $\Sigma$ -tree  $\tau$  is given by a nonempty subset  $\text{Dom}(\tau) \subseteq \mathbb{N}^*$  (called the *domain* of  $\tau$ ) and a *labeling* function  $\tau: \text{Dom}(\tau) \rightarrow \Sigma$  that are subject to the following conditions.<sup>3</sup>

- (1)  $\text{Dom}(\tau)$  is *prefix-closed*: for any  $w \in \mathbb{N}^*$  and  $i \in \mathbb{N}$ ,  $wi \in \text{Dom}(\tau)$  implies  $w \in \text{Dom}(\tau)$ . See Figure 1.
- (2)  $\text{Dom}(\tau)$  is *lower-closed*: for any  $w \in \mathbb{N}^*$  and  $i, j \in \mathbb{N}$ ,  $wj \in \text{Dom}(\tau)$  and  $i \leq j$  imply  $wi \in \text{Dom}(\tau)$ . See Figure 1.
- (3) The labeling function is consistent with arities: for any  $w \in \text{Dom}(\tau)$ , let  $\sigma = \tau(w)$ . Then  $w0, w1, \dots, w(|\sigma| - 1)$  belong to  $\text{Dom}(\tau)$ , and  $wi \notin \text{Dom}(\tau)$  for each  $i$  such that  $|\sigma| \leq i$ . See Figure 2.

The set of all  $\Sigma$ -trees shall be denoted by  $\text{Tree}_\Sigma$ .

Intuitively, a  $\Sigma$ -tree is a possibly infinite tree whose nodes are labeled from  $\Sigma$  and each node, say labeled by  $\sigma$ , has precisely  $|\sigma|$  children. A sequence  $w \in \mathbb{N}^*$  is understood as a *position* in a tree.

The following definitions are standard, too, in the tree-automata literature.

**Definition 3.5** (run). A *run*  $\rho$  of an NBTA (Definition 3.1)  $\mathcal{X} = (X, \Sigma, \delta, I, \text{Acc})$  is a (possibly infinite) tree whose nodes are  $(\Sigma \times X)$ -labeled. That should be consistent with arities of symbols, and compatible with the initial states ( $I \subseteq X$ ) and the transition  $\delta$  of the automaton  $\mathcal{X}$ . Precisely, it is given by the following conditions:

- (1) A nonempty subset  $\text{Dom}(\rho) \subseteq \mathbb{N}^*$  that is subject to the same conditions (of being prefix-closed and lower-closed) as for  $\Sigma$ -trees (Definition 3.4).
- (2) A labeling function  $\rho: \text{Dom}(\rho) \rightarrow \Sigma \times X$  such that, if  $\rho(w) = (\sigma, x)$ , then  $w$  has precisely  $|\sigma|$  successors  $w0, w1, \dots, w(|\sigma| - 1) \in \text{Dom}(\rho)$ .
- (3) Successors are reachable by a transition, in the sense that  $(\sigma_w, (x_{w0}, \dots, x_{w(|\sigma|-1)})) \in \delta(x_w)$  holds, where  $\rho(w)$  is labeled with  $(\sigma_w, x_w)$ , and  $\rho(wi)$  is labeled with  $(\sigma_{wi}, x_{wi})$  for each  $0 \leq i < |\sigma|$ .
- (4) The root is labeled with an initial state, that is,  $x_\varepsilon \in I$  where  $\rho(\varepsilon) = (\sigma_\varepsilon, x_\varepsilon)$ .

The set of all runs of the NBTA  $\mathcal{X}$  is denoted by  $\text{Run}_\mathcal{X}^\rho$ .

<sup>3</sup>We shall use the same notation  $\tau$  for a tree itself and its labeling function. Confusion is unlikely.

The map denoted by  $\text{DelSt}: \text{Run}_{\mathcal{X}}^{\mathcal{P}} \rightarrow \text{Tree}_{\Sigma}$  takes a run  $\rho \in \text{Run}_{\mathcal{X}}^{\mathcal{P}}$ , removes its  $X$ -labels applying the first projection to each label, and returns the resulting  $\Sigma$ -labeled tree. The resulting tree is easily seen to be a  $\Sigma$ -tree by Definition 3.4. We say that a run  $\rho$  is *over* the  $\Sigma$ -tree  $\text{DelSt}(\rho)$ .

In summary, a (possibly infinite)  $(\Sigma \times X)$ -labeled tree  $\rho$  is a run of an NBTA  $\mathcal{X} = (X, \Sigma, \delta, I, \text{Acc})$  if: the  $X$ -label of its root is initial  $s \in I$ ; and for each node with a label  $(\sigma, x)$ , it has  $|\sigma|$  children and we have  $(\sigma, (x_1, \dots, x_{|\sigma|})) \in \delta(x)$  where  $x_1, \dots, x_{|\sigma|}$  are the  $X$ -labels of its children.

We next define a notion of branch.

**Definition 3.6** (branch). Let  $\tau$  be a  $\Sigma$ -tree. A *branch* of  $\tau$  is either:

- an infinite sequence  $\pi = \pi_1\pi_2\dots \in \mathbb{N}^{\omega}$  (where  $\pi_i \in \mathbb{N}$ ) such that any finite prefix  $\pi_{\leq n} = \pi_1\dots\pi_n$  of it belongs to  $\text{Dom}(\tau)$ ; or
- a finite sequence  $\pi = \pi_1\dots\pi_n \in \mathbb{N}^*$  where  $\pi_i \in \mathbb{N}$  that belongs to  $\text{Dom}(\tau)$  and such that  $\pi 0 \notin \text{Dom}(\tau)$ .<sup>4</sup>

The set of all branches of a  $\Sigma$ -tree  $\tau$  is denoted by  $\text{Branch}(\tau)$ . The notion of branch is defined similarly for a run, with  $\text{Branch}(\rho)$  denoting the set of all branches of  $\rho$ .

We define a notion of accepting run. A run  $\rho$  of an NBTA  $\mathcal{X}$  is said to be *accepting* if any infinite branch  $\pi$  of the tree  $\rho$  satisfies the Büchi acceptance condition, that is, it visits accepting states (in  $\text{Acc}$ ) infinitely often. The sets of runs and accepting runs of  $\mathcal{X}$  are denoted by  $\text{Run}_{\mathcal{X}}^{\mathcal{P}}$  and  $\text{AccRun}_{\mathcal{X}}^{\mathcal{P}}$ , respectively. Formally, they are defined as follows.

**Definition 3.7** (accepting run). A run  $\rho$  of an NBTA  $\mathcal{X} = (X, \Sigma, \delta, I, \text{Acc})$  is said to be *accepting* if, any branch  $\pi \in \text{Branch}(\rho)$  of it is *accepting* in the following sense.

- The branch  $\pi$  is an infinite sequence  $\pi = \pi_1\pi_2\dots \in \mathbb{N}^{\omega}$ , and the labels along the branch  $(\sigma_{\varepsilon}, x_{\varepsilon})(\sigma_{\pi_1}, x_{\pi_1})(\sigma_{\pi_1\pi_2}, x_{\pi_1\pi_2})\dots$  (here  $(\sigma_w, x_w) := \rho(w)$  for each  $w \in \mathbb{N}^*$ ) visit accepting states infinitely often, that is, there exists an infinite sequence  $n_1 < n_2 < \dots$  of natural numbers such that  $x_{\pi_1\dots\pi_{n_i}} \in F$  for each  $i \in \mathbb{N}$ ; or
- the branch  $\pi$  is a finite sequence  $\pi = \pi_1\dots\pi_n \in \mathbb{N}^*$ .

The set of all accepting runs over  $\mathcal{X}$  is denoted by  $\text{AccRun}_{\mathcal{X}}^{\mathcal{P}}$ .

Using the notions defined so far, we can define accepted languages of NBTAs as follows.

**Definition 3.8** (accepted language  $L(\mathcal{X})$ ). For an NBTA  $\mathcal{X}$ , its (*Büchi*) *language*  $L(\mathcal{X}) \subseteq \text{Tree}_{\Sigma}$  is defined by  $L(\mathcal{X}) = \{\text{DelSt}(\rho) \mid \rho \in \text{AccRun}_{\mathcal{X}}^{\mathcal{P}}\}$ .

**Example 3.9.** For the NBTA  $\mathcal{X}$  in Example 3.2, the Büchi language  $L(\mathcal{X})$  collects all the  $\{a, b\}$ -labeled infinite binary trees where  $b$  appears infinitely many times on each branch.

**3.2. Fair Simulation for Nondeterministic Büchi Tree Automata.** In this section we introduce *fair simulation* for NBTAs; this is our first contribution. For finite-state NBTAs, our fair simulation notion is essentially the same as the one in [vB08]. However, unlike the notion in [vB08] that is defined combinatorially via a parity game, ours is expressed by means of equational systems (Section 2), hence is applicable to infinitary settings.

Here is a brief description of a parity game. For formal definitions, see [TW<sup>+</sup>02] for example. A *parity game* is a game played by two players called *Even* and *Odd* over a

<sup>4</sup>This means that  $\pi$  is a leaf of  $\tau$ , and that  $\tau(\pi)$  is a 0-ary symbol.

finite-state directed graph  $G = (V, E)$ . Each node  $v \in V$  is called a *position*, and the set  $V$  of positions is divided into two parts—the one where Even chooses the next move and the one where Odd chooses the next move. We assume that a game is equipped with a *priority function*  $p : V \rightarrow \{0, 1, \dots, n\}$  that assigns a natural number called a *priority* to each state.

Once an initial state  $v_0$  and strategies (functions from finite sequences of positions to a position) for Even and Odd are fixed, a run  $\rho = v_0 v_1 \dots \in V^\omega$ , an infinite sequence over  $G$ , is determined in a natural manner. A run is *winning* for Even (respectively Odd) if the *maximum* priority that appears infinitely often in  $\rho$  is even (respectively odd). A parity game is said to be *winning* for Even from a position  $v_0$  if there exists a strategy for Even such that, regardless of the strategy of Odd, the resulting run from  $v_0$  is winning for Even. A notion of winning for Odd is defined similarly. It is known that parity games satisfy *determinacy* [TW<sup>+</sup>02]: for each parity game and each state in it, the game is winning from the state for exactly one of Even and Odd.

We hereby review the combinatorial definition of fair simulation in [vB08] via a parity game, to show an intuition behind our definition.

**Definition 3.10** (parity game for NBTA fair simulation, [vB08]). Let  $\mathcal{X} = (X, \Sigma, \delta_{\mathcal{X}}, I_{\mathcal{X}}, \text{Acc}_{\mathcal{X}})$  and  $\mathcal{Y} = (Y, \Sigma, \delta_{\mathcal{Y}}, I_{\mathcal{Y}}, \text{Acc}_{\mathcal{Y}})$  be NBTAs such that  $X$  and  $Y$  are finite. Let  $X_1 = X \setminus \text{Acc}_{\mathcal{X}}$ ,  $X_2 = \text{Acc}_{\mathcal{X}}$ , and similarly for  $Y = Y_1 \cup Y_2$ . We define a parity game  $G_{\mathcal{X}, \mathcal{Y}}$  as follows.

Position	Player	The set of possible moves	Priority
*	Odd	$I_{\mathcal{X}}$	0
$x \in X$	Even	$\{(x, y) \mid y \in I_{\mathcal{Y}}\}$	0
$(x, y) \in X \times Y$	Odd	$\left\{ \left( \begin{array}{c} (\sigma, x_1, \dots, x_{ \sigma }) \\ y \end{array} \right) \mid \begin{array}{c} (\sigma, x_1, \dots, x_{ \sigma }) \\ \in \delta_{\mathcal{X}}(x) \end{array} \right\}$	$\begin{cases} 0 & ((x, y) \in X_1 \times Y_1) \\ 1 & ((x, y) \in X_2 \times Y_1) \\ 2 & ((x, y) \in X \times Y_2) \end{cases}$
$((\sigma, x_1, \dots, x_{ \sigma }), y)$ $\in (\prod_{\sigma \in \Sigma} X^{ \sigma }) \times Y$	Even	$\left\{ \left( \begin{array}{c} (x_1, y_1), \dots, \\ (x_{ \sigma }, y_{ \sigma }) \end{array} \right) \mid \begin{array}{c} (\sigma, y_1, \dots, y_{ \sigma }) \\ \in \delta_{\mathcal{Y}}(y) \end{array} \right\}$	0
$(p_1, \dots, p_n) \in (X \times Y)^*$	Odd	$\{p_i \mid 1 \leq i \leq n\}$	0

Note that as the number of positions of the game is finite, the problem to determine the winner of  $G_{\mathcal{X}, \mathcal{Y}}$  is decidable.

We now introduce our fair simulation notion by means of equational systems. We will later show that for finite-state NBTAs, our simulation notion is essentially the same as the one in Definition 3.10.

**Definition 3.11** (fair simulation for NBTAs). Let  $\mathcal{X} = (X, \Sigma, \delta_{\mathcal{X}}, I_{\mathcal{X}}, \text{Acc}_{\mathcal{X}})$  and  $\mathcal{Y} = (Y, \Sigma, \delta_{\mathcal{Y}}, I_{\mathcal{Y}}, \text{Acc}_{\mathcal{Y}})$  be NBTAs. We define  $X_1, X_2$  and  $Y_1, Y_2$  as in Definition 3.10. A *fair simulation* from  $\mathcal{X}$  to  $\mathcal{Y}$  is a relation  $R \subseteq X \times Y$  such that:

- (1) For all  $x \in I_{\mathcal{X}}$ , there exists  $y \in I_{\mathcal{Y}}$  such that  $(x, y) \in R$ .
- (2) Let  $u_1^{\text{sol}}, \dots, u_4^{\text{sol}}$  be the solution of the following equational system (note  $\mu$ 's vs.  $\nu$ 's).

$$\begin{aligned}
 u_1 &=_{\nu} \Box_{\mathcal{X}, 1}(\Diamond_{\mathcal{Y}, 1}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) && \subseteq X_1 \times Y_1 \\
 u_2 &=_{\mu} \Box_{\mathcal{X}, 2}(\Diamond_{\mathcal{Y}, 1}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) && \subseteq X_2 \times Y_1 \\
 u_3 &=_{\nu} \Box_{\mathcal{X}, 1}(\Diamond_{\mathcal{Y}, 2}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) && \subseteq X_1 \times Y_2 \\
 u_4 &=_{\nu} \Box_{\mathcal{X}, 2}(\Diamond_{\mathcal{Y}, 2}(\bigwedge_{\Sigma}(u_1 \cup u_2 \cup u_3 \cup u_4))) && \subseteq X_2 \times Y_2
 \end{aligned} \tag{3.1}$$

Then  $R$  is below the solution, that is,  $R \subseteq u_1^{\text{sol}} \cup \dots \cup u_4^{\text{sol}}$ .

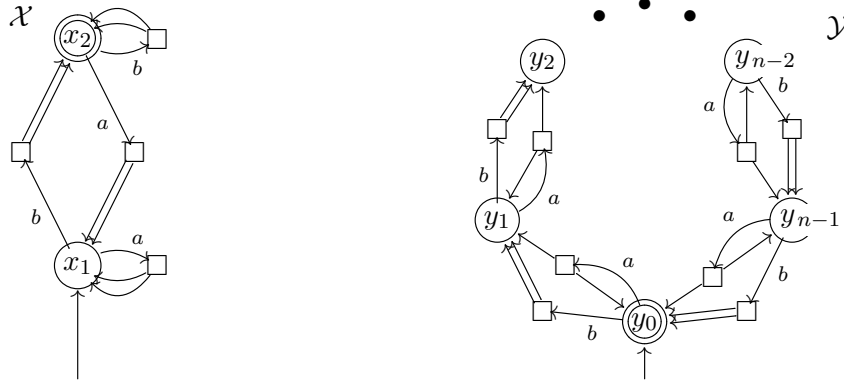
Here the functions  $\square_{\mathcal{X},i} : \mathcal{P}(\prod_{\sigma \in \Sigma} X^{|\sigma|} \times Y) \rightarrow \mathcal{P}(X_i \times Y)$ ,  $\diamond_{\mathcal{Y},j} : \mathcal{P}(\prod_{\sigma \in \Sigma} X^{|\sigma|} \times \prod_{\sigma \in \Sigma} Y^{|\sigma|}) \rightarrow \mathcal{P}(\prod_{\sigma \in \Sigma} X^{|\sigma|} \times Y_j)$  and  $\bigwedge_{\Sigma} : \mathcal{P}(X \times Y) \rightarrow \mathcal{P}(\prod_{\sigma \in \Sigma} X^{|\sigma|} \times \prod_{\sigma \in \Sigma} Y^{|\sigma|})$  are defined as follows.

$$\begin{aligned} \square_{\mathcal{X},i}(S) &:= \{(x, y) \in X_i \times Y \mid \forall \mathbf{x}' \in \delta_{\mathcal{X}}(x). (\mathbf{x}', y) \in S\} \\ \diamond_{\mathcal{Y},j}(T) &:= \{(\mathbf{x}', y) \in \prod_{\sigma \in \Sigma} X^{|\sigma|} \times Y_j \mid \exists \mathbf{y}' \in \delta_{\mathcal{Y}}(y). (\mathbf{x}', \mathbf{y}') \in T\} \\ \bigwedge_{\Sigma}(U) &:= \left\{ \begin{array}{l} ((\sigma, x_1, \dots, x_{|\sigma|}), (\sigma', y_1, \dots, y_{|\sigma'|})) \\ \in \prod_{\sigma \in \Sigma} X^{|\sigma|} \times \prod_{\sigma \in \Sigma} Y^{|\sigma|} \mid \sigma = \sigma', \\ \forall i. (x_i, y_i) \in U \end{array} \right\} \end{aligned}$$

**Theorem 3.12** (soundness). *In the setting of Definition 3.11, existence of a fair simulation from  $\mathcal{X}$  to  $\mathcal{Y}$  implies language inclusion, that is,  $L(\mathcal{X}) \subseteq L(\mathcal{Y})$ .*

Our proof of Theorem 3.12 relies on a categorical theory developed in later sections, and will be given in Section 6.3.

**Example 3.13.** Let  $\mathcal{X}$  and  $\mathcal{Y}$  be the NBTA's illustrated below, where a transition  $z \xrightarrow{\sigma} (z_1, z_2)$  is represented by  $z \xrightarrow{\sigma} \square \rightrightarrows \begin{matrix} z_1 \\ z_2 \end{matrix}$ .



Here the ranked alphabet is given by  $\Sigma = \{a, b\}$  where  $|a| = |b| = 2$ . Let  $X$  and  $Y$  be the state spaces of  $\mathcal{X}$  and  $\mathcal{Y}$  respectively, and define  $X_1, X_2$  and  $Y_1, Y_2$  as in Definition 3.11.

We can see that  $R_1 = X_1 \times Y_1$ ,  $R_2 = X_2 \times Y_1$ ,  $R_3 = X_1 \times Y_2$  and  $R_4 = X_2 \times Y_2$  are the solution of the equational system (3.1) in Definition 3.11 induced by  $\mathcal{X}$  and  $\mathcal{Y}$  here. Hence  $R = X \times Y$  is a fair simulation from  $\mathcal{X}$  to  $\mathcal{Y}$ , and this implies language inclusion.

We conclude this section by showing a relationship between the simulation notion via parity games (Definition 3.10) and our simulation notion (Definition 3.11). Roughly speaking, a parity game is understood as a combinatorial presentation of an equational system like (3.1) over finite lattices  $L_1, \dots, L_m$  [HSC16]. If NBTA's  $\mathcal{X}$  and  $\mathcal{Y}$  have finite state-spaces, translating (3.1) leads to the parity game in Definition 3.10. Formally, we have the following proposition. The proof is similar to the one for [HSC15, Corollary A.5].

**Proposition 3.14.** *Let  $\mathcal{X} = (X, \Sigma, \delta_{\mathcal{X}}, I_{\mathcal{X}}, \text{Acc}_{\mathcal{X}})$  and  $\mathcal{Y} = (Y, \Sigma, \delta_{\mathcal{Y}}, I_{\mathcal{Y}}, \text{Acc}_{\mathcal{Y}})$  be NBTA's such that  $X$  and  $Y$  are finite. Then a fair simulation (Def. 3.11) from  $\mathcal{X}$  to  $\mathcal{Y}$  exists if and only if the player Even is winning in the parity game  $G_{\mathcal{X}, \mathcal{Y}}$  in Def. 3.10 from  $*$ ; if that is the case we have  $L(\mathcal{X}) \subseteq L(\mathcal{Y})$ .  $\square$*

## 4. FAIR SIMULATION FOR FINITE-STATE PROBABILISTIC BÜCHI WORD AUTOMATA

This is the second section in which we describe our technical contributions in concrete set-theoretic terms. They are derived from the theoretical backgrounds that we describe in later sections. In this section we focus on probabilistic systems.

In what follows we adopt the following conventions. The  $(x, y)$ -entry of a matrix  $A \in [0, 1]^{X \times Y}$  is denoted by  $A_{x,y}$ ; the  $x$ -th entry of a vector  $\iota \in [0, 1]^X$  is  $\iota_x$ . For  $A, B \in [0, 1]^{X \times Y}$ , we write  $A \leq B$  if  $A_{x,y} \leq B_{x,y}$  for all  $x$  and  $y$ .

**Definition 4.1** (PBWA). A (*generative*) *probabilistic Büchi word automaton* (PBWA) is a quintuple  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$  consisting of a countable state space  $X$ , a countable alphabet  $\mathbf{A}$ , transition matrices  $M(a) \in [0, 1]^{X \times X}$  for each  $a \in \mathbf{A}$ , an initial distribution  $\iota \in [0, 1]^X$ , and a set  $\text{Acc} \subseteq X$  of accepting states. We require that the matrices  $M(a)$  and the vector  $\iota$  are substochastic:  $\sum_{a \in \mathbf{A}} \sum_{x' \in X} (M(a))_{x,x'} \leq 1$  for each  $x \in X$ , and  $\sum_{x \in X} \iota_x \leq 1$ .

Note that the initial vector and transition matrices are *sub*-stochastic:  $\sum_{a \in \mathbf{A}} \sum_{x' \in X} (M(a))_{x,x'}$  and  $\sum_{x \in X} \iota_x$  are allowed to be strictly smaller than 1. The missing probabilities are for *divergence*. We require  $\sum_a \sum_{x'} (M(a))_{x,x'} \leq 1$ : this means our automaton is *generative* and it chooses which character  $a \in \mathbf{A}$  to *output*. This is in contrast to a *reactive* automaton (that *reads* characters), in which case we would require  $\sum_{x'} (M(a))_{x,x'} \leq 1$  for each  $a$ .

**Example 4.2.** We define a PBWA  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$  as follows.

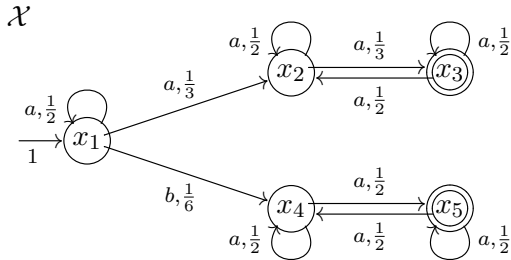
- $X = \{x_1, x_2, x_3, x_4, x_5\}$
- $\mathbf{A} = \{a, b\}$

$$\bullet M(a) = \begin{matrix} & \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{matrix} & \begin{pmatrix} 1/2 & 1/3 & 0 & 0 & 0 \\ 0 & 1/2 & 1/3 & 0 & 0 \\ 0 & 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 1/2 & 1/2 \end{pmatrix} \end{matrix} \quad \text{and} \quad M(b) = \begin{matrix} & \begin{matrix} x_1 & x_2 & x_3 & x_4 & x_5 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1/6 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

$$\bullet \iota = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

- $\text{Acc} = \{x_3, x_5\}$

Then  $\mathcal{X}$  is illustrated as below.



In the next section we shall give a definition of accepted languages of PBWAs. This is rather standard (see [CHS14] for a reactive variant).

**4.1. Accepted Languages of Probabilistic Büchi Word Automata.** The language  $L(\mathcal{X})$ —a subprobability measure that tells which words are generated by what probabilities—is essentially the *push-forward measure* [Doo94] obtained from the one over the set  $\text{Run}_{\mathcal{X}}^{\mathcal{G}}$  of *runs* of a PBWA.

**Definition 4.3** (run). For a PBWA  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$ , a *run* over  $\mathcal{X}$  is an infinite word  $\rho \in (\mathbf{A} \times X)^\omega$ . The set of all runs over  $\mathcal{X}$  is denoted by  $\text{Run}_{\mathcal{X}}^{\mathcal{G}}$ . A *partial run* over  $\mathcal{X}$  is a finite word  $\xi \in (\mathbf{A} \times X)^* \times X$ . A run  $\rho = (a_0, x_0)(a_1, x_1) \dots \in \text{Run}_{\mathcal{X}}^{\mathcal{G}}$  is *accepting* if  $x_i \in \text{Acc}$  for infinitely many  $i$ 's. The set of all accepting runs over  $\mathcal{X}$  is denoted by  $\text{AccRun}_{\mathcal{X}}^{\mathcal{G}}$ .

We define the language of  $\mathcal{X}$  as a subprobability measure over the set  $\mathbf{A}^\omega$  of infinite words. The set  $\mathbf{A}^\omega$  of all infinite words over  $\mathbf{A}$  carries a canonical ‘‘cylindrical’’ measurable structure generated by  $\{w\mathbf{A}^\omega \mid w \in \mathbf{A}^*\}$  (see [BK08] for example). The set  $(\mathbf{A} \times X)^\omega$  of runs comes with a cylindrical measurable structure, too.

**Definition 4.4.** Let  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$  be a PBWA. For  $w \in \mathbf{A}^*$ , the *cylinder set* generated by  $w$  is a set

$$\text{Cyl}(w) := \{ww' \in \mathbf{A}^\omega \mid w \in \mathbf{A}^*, w' \in \mathbf{A}^\omega\}.$$

We write  $\mathfrak{F}_{\mathbf{A}^\omega}$  for the smallest  $\sigma$ -algebra over  $\mathbf{A}^\omega$  that is generated by the cylinder sets  $\{\text{Cyl}(w) \mid w \in \mathbf{A}^*\}$ .

Similarly, for a partial run  $\xi = (a_0, x_0) \dots (a_{i-1}, x_{i-1})x_i \in (\mathbf{A} \times X)^* \times X$ , the *cylinder set* generated by  $\xi$  is a set

$$\text{Cyl}_{\mathcal{X}}(\xi) := \{(a_0, x_0) \dots (a_i, x_i)(a_{i+1}, x_{i+1}) \dots \in \text{Run}_{\mathcal{X}}^{\mathcal{G}} \mid a_i, a_{i+1}, \dots \in \mathbf{A}, x_{i+1}, x_{i+2}, \dots \in X\}.$$

We write  $\mathfrak{F}_{\mathcal{X}}$  for the  $\sigma$ -algebra over  $\text{Run}_{\mathcal{X}}^{\mathcal{G}}$  generated by the cylinder sets  $\{\text{Cyl}_{\mathcal{X}}(\xi) \mid \xi \in (\mathbf{A} \times X)^* \times X\}$ .

We define  $\text{DelSt} : \text{Run}_{\mathcal{X}}^{\mathcal{G}} \rightarrow \mathbf{A}^\omega$  by  $\text{DelSt}((a_0, x_0)(a_1, x_1) \dots) := a_0a_1 \dots$ .

Now it can be shown that the set  $\text{AccRun}_{\mathcal{X}}^{\mathcal{G}}$  of *accepting* runs—visiting  $\odot$  infinitely often—is a measurable subset. This result (as stated in the following lemma) is much like [CHS14, Lemma 36] and hardly novel.

**Lemma 4.5.** *The set  $\text{AccRun}_{\mathcal{X}}^{\mathcal{G}}$  of accepting runs is an  $\mathfrak{F}_{\mathcal{X}}$ -measurable subset of  $\text{Run}_{\mathcal{X}}^{\mathcal{G}}$ .*

*Proof.* For each  $k \in \mathbb{N}$ , we define a set  $\text{NAccRun}_k^{\text{inf}} \subseteq \text{Run}_{\mathcal{X}}$  as follows.

$$\text{NAccRun}_k^{\text{inf}} := \{(a_0, x_0)(a_1, x_1) \dots \mid x_k \notin \text{Acc}\}. \quad (4.1)$$

Then we have:

$$\text{NAccRun}_k^{\text{inf}} = \bigcup_{a_0 \in \mathbf{A}} \dots \bigcup_{a_{k-1} \in \mathbf{A}} \bigcup_{x_0 \in X} \dots \bigcup_{x_{k-1} \in X} \bigcup_{x_k \in X \setminus \text{Acc}} \text{Cyl}((a_0, x_0) \dots (a_{k-1}, x_{k-1})x_k).$$

As  $X$  and  $\mathbf{A}$  are countable sets, by definition of the  $\sigma$ -algebra  $\mathfrak{F}_{\mathcal{X}}$ ,  $\text{NAccRun}_k^{\text{inf}}$  is measurable.

By definition of  $\text{AccRun}_{\mathcal{X}}$ , it is easy to see that:

$$\text{AccRun}_{\mathcal{X}} = \text{Run}_{\mathcal{X}} \setminus \bigcup_{m \in \mathbb{N}} \bigcap_{n \in \mathbb{N}} \text{NAccRun}_{m+n}^{\text{inf}}.$$

Hence  $\text{AccRun}_{\mathcal{X}}$  is measurable. □

The following notion of *no-divergence* probability plays an important role. Recall that a PBWA can exhibit divergence.

**Definition 4.6** ( $\text{NoDiv}_{\mathcal{X}}$ ). Let  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$  be a PBWA. For each  $k \in \mathbb{N}$ ,  $\text{NoDiv}_{\mathcal{X},k} : X \rightarrow [0, 1]$  is defined inductively by:

$$\begin{aligned} \text{NoDiv}_{\mathcal{X},0}(x) &:= 1 \text{ ,} \\ \text{NoDiv}_{\mathcal{X},k+1}(x) &:= \sum_{a \in \mathbf{A}} \sum_{x' \in X} (M(a))_{x,x'} \cdot \text{NoDiv}_{\mathcal{X},k}(x') \text{ .} \end{aligned} \quad (4.2)$$

Note that as  $\sum_{a \in \mathbf{A}} \sum_{x' \in X} (M(a))_{x,x'} \leq 1$  for each  $x$ ,  $\text{NoDiv}_{\mathcal{X},k}(x)$  is decreasing with respect to  $k$ . We define a function  $\text{NoDiv}_{\mathcal{X}} : X \rightarrow [0, 1]$  by  $\text{NoDiv}_{\mathcal{X}}(x) := \lim_{k \rightarrow \infty} \text{NoDiv}_{\mathcal{X},k}(x)$ .

Intuitively,  $\text{NoDiv}_{\mathcal{X}}(x)$  is a probability in which an execution of  $\mathcal{X}$  from the state  $x$  does not exhibit divergence. This probability is used to define probabilistic accepted languages of PBWAs.

We can now define a subprobability measure  $\mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}$  on  $\text{Run}_{\mathcal{X}}^{\mathcal{G}}$  induced by the PBWA  $\mathcal{X}$  (via the Carathéodory theorem).

**Definition 4.7** ( $\mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}$  over  $\text{Run}_{\mathcal{X}}^{\mathcal{G}}$ ). Let  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$  be a PBWA. We shall define a subprobability measure  $\mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}$  over  $(\text{Run}_{\mathcal{X}}^{\mathcal{G}}, \mathfrak{F}_{\mathcal{X}})$ . It is given, for each partial run  $\xi = (a_0, x_0) \dots (a_{i-1}, x_{i-1})x_i \in (\mathbf{A} \times X)^* \times X$ , by

$$\mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}(\text{Cyl}_{\mathcal{X}}(\xi)) := \iota_{x_0} \cdot P_{\mathcal{X}}(\xi) \text{ .} \quad (4.3)$$

Here  $P_{\mathcal{X}}(\xi)$  is defined inductively as follows.

$$P_{\mathcal{X}}(\xi) := \begin{cases} \text{NoDiv}_{\mathcal{X}}(x_0) & (i = 0) \\ (M(a_0))_{x_0,x_1} \cdot P_{\mathcal{X}}((a_1, x_1) \dots (a_{i-1}, x_{i-1})x_i) & (i > 0) \end{cases}$$

**Proposition 4.8.** *Definition 4.7 is well-defined. That is to say, there exists a unique subprobability measure  $\mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}$  over  $(\text{Run}_{\mathcal{X}}^{\mathcal{G}}, \mathfrak{F}_{\mathcal{X}})$  that satisfies the equation (4.3).*

*Proof.* We first prove that for each  $i \in \mathbb{N}$ ,  $a_0, \dots, a_{i-1} \in \mathbf{A}$  and  $x_0, \dots, x_i \in X$ , we have:

$$P_{\mathcal{X}}((a_0, x_0) \dots (a_{i-1}, x_{i-1})x_i) = \sum_{a_i \in \mathbf{A}} \sum_{x_{i+1} \in X} P_{\mathcal{X}}((a_0, x_0) \dots (a_{i-1}, x_{i-1})(a_i, x_i)x_{i+1}) \text{ .}$$

We prove it by the induction on  $i$ .

- If  $i = 0$ , then  $\xi = x_0$  and hence we have:

$$\begin{aligned} P_{\mathcal{X}}(\xi) &= \text{NoDiv}_{\mathcal{X}}(x_0) \\ &= \lim_{k \rightarrow \infty} \text{NoDiv}_{\mathcal{X},k}(x) \\ &= \lim_{k \rightarrow \infty} \sum_{a \in \mathbf{A}} \sum_{x_1 \in X} (M(a))_{x_0,x_1} \cdot \text{NoDiv}_{\mathcal{X},k-1}(x_1) \\ &= \lim_{k \rightarrow \infty} \sum_{a \in \mathbf{A}} \sum_{x_1 \in X} (M(a))_{x_0,x_1} \cdot \text{NoDiv}_{\mathcal{X},k}(x_1) \\ &= \sum_{a_0 \in \mathbf{A}} \sum_{x_1 \in X} (M(a_0))_{x_0,x_1} \cdot \left( \lim_{k \rightarrow \infty} \text{NoDiv}_{\mathcal{X},k}(x_1) \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{a_0 \in \mathbf{A}} \sum_{x_1 \in X} (M(a_0))_{x_0, x_1} \cdot P_{\mathcal{X}}(x_1) \\
&= \sum_{a_0 \in \mathbf{A}} \sum_{x_1 \in X} P_{\mathcal{X}}((a_0, x_0)x_1)
\end{aligned}$$

- If  $i > 0$ , then we have:

$$\begin{aligned}
P_{\mathcal{X}}(\xi) &= (M(a_0))_{x_0, x_1} \cdot P_{\mathcal{X}}((a_1, x_1) \dots (a_{i-1}, x_{i-1})x_i) \\
&= (M(a_0))_{x_0, x_1} \cdot \sum_{a_i \in \mathbf{A}} \sum_{x_{i+1} \in X} P_{\mathcal{X}}((a_1, x_1) \dots (a_{i-1}, x_{i-1})(a_i, x_i)x_{i+1}) \\
&\hspace{25em} \text{(by the induction hypothesis)} \\
&= \sum_{a_i \in \mathbf{A}} \sum_{x_{i+1} \in X} (M(a_0))_{x_0, x_1} \cdot P_{\mathcal{X}}((a_1, x_1) \dots (a_{i-1}, x_{i-1})(a_i, x_i)x_{i+1}) \\
&= \sum_{a_i \in \mathbf{A}} \sum_{x_{i+1} \in X} P_{\mathcal{X}}((a_0, x_0)(a_1, x_1) \dots (a_{i-1}, x_{i-1})(a_i, x_i)x_{i+1}).
\end{aligned}$$

Hence we have:

$$\mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}((a_0, x_0) \dots (a_{i-1}, x_{i-1})x_i) = \sum_{a_i \in \mathbf{A}} \sum_{x_{i+1} \in X} \mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}((a_0, x_0) \dots (a_{i-1}, x_{i-1})(a_i, x_i)x_{i+1}).$$

Therefore Proposition 4.8 is immediate from Carathéodory's extension theorem (see [ADD00] for example).  $\square$

Now we can define the language of a PBWA  $\mathcal{X}$ .

**Definition 4.9** (language of PBWA).  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$  be a PBWA. A subprobability measure  $L(\mathcal{X})$  over  $(\mathbf{A}^\omega, \mathfrak{F}_{\mathbf{A}^\omega})$  is defined as follows: for each  $w \in \mathbf{A}^*$ ,

$$L(\mathcal{X})(\text{Cyl}(w)) := \mu_{\mathcal{X}}^{\text{Run}_{\mathcal{X}}^{\mathcal{G}}}(\text{DelSt}^{-1}(\text{Cyl}(w)) \cap \text{AccRun}_{\mathcal{X}}^{\mathcal{G}}) . \quad (4.4)$$

Note here that  $\text{DelSt}^{-1}(\text{Cyl}(w)) = \bigcup_{\xi \in \text{DelSt}^{-1}(\{w\})} \text{Cyl}_{\mathcal{X}}(\xi)$ . As  $w$  is a finite word and the state space  $X$  is countable, the union in the above equation is a countable one. Hence the set  $\text{DelSt}^{-1}(\text{Cyl}(w))$  is measurable.

The following proposition can be proved in a similar manner to Proposition 4.8.

**Proposition 4.10.** *Definition 4.9 is well-defined. That is, there exists a unique subprobability measure  $L(\mathcal{X})$  over  $(\Sigma^\omega, \mathfrak{F}_{\mathbf{A}^\omega})$  that satisfies the equation (4.4).*  $\square$

**Example 4.11.** Let  $\mathcal{X}$  be the PBWA in Example 4.2. For each cylinder set  $w\mathbf{A}^\omega$  where  $w \in \mathbf{A}^*$ , the subprobability measure  $L(\mathcal{X})$  assigns the following probability.

$$L(\mathcal{X})(w\mathbf{A}^\omega) = \begin{cases} \frac{1}{2^n} \cdot \frac{1}{3} & (w = \underbrace{a \dots a}_n, \underbrace{a \dots a b}_n \text{ or } \underbrace{a \dots a b a \dots a}_n) \\ 0 & \text{(otherwise)} \end{cases}$$



**4.2. Fair Simulation for PBWAs.** We continue to introduce *fair simulation* for PBWAs. This is one of our main contributions: to the best of our knowledge this is the first one for *probabilistic* Büchi (word) automata. Note that our simulation is given by a matrix and not by a relation; this follows our previous work [Has06, UH17].

**Definition 4.12** (fair simulation for PBWAs). Let  $\mathcal{X} = (X, \mathbf{A}, M_{\mathcal{X}}, \iota_{\mathcal{X}}, \text{Acc}_{\mathcal{X}})$  and  $\mathcal{Y} = (Y, \mathbf{A}, M_{\mathcal{Y}}, \iota_{\mathcal{Y}}, \text{Acc}_{\mathcal{Y}})$  be probabilistic Büchi word automata with the same alphabet  $\mathbf{A}$ . Let  $A$  be a matrix such that  $A \in [0, 1]^{Y \times X}$ . We define  $X_1 = X \setminus \text{Acc}_{\mathcal{X}}$  and  $X_2 = \text{Acc}_{\mathcal{X}}$  (like in Definition 3.11), and similarly for  $Y_1$  and  $Y_2$ . Moreover, let  $M_{\mathcal{X},i}(a) \in [0, 1]^{X_i \times X}$ ,  $M_{\mathcal{Y},j}(a) \in [0, 1]^{Y_j \times Y}$  and  $A_{ji} \in [0, 1]^{Y_j \times X_i}$  denote the obvious partial matrices of  $M_{\mathcal{X}}(a) \in [0, 1]^{X \times X}$ ,  $M_{\mathcal{Y}}(a) \in [0, 1]^{Y \times Y}$  and  $A \in [0, 1]^{Y \times X}$ , respectively. We say that the matrix  $A$  is a *fair simulation* from  $\mathcal{X}$  to  $\mathcal{Y}$  if it satisfies the following conditions.

- (1) The matrix  $A$  is a substochastic matrix:  $\sum_{x \in X} A_{y,x} \leq 1$  for each  $y \in Y$ .
- (2) The matrix  $A$  is a *forward simulation matrix* [UH14, UH17], that is,  $\iota_{\mathcal{X}} \leq \iota_{\mathcal{Y}} \cdot A$  and  $A \cdot M_{\mathcal{X}}(a) \leq M_{\mathcal{Y}}(a) \cdot A$  for each  $a \in \mathbf{A}$ .
- (3) The partial matrices  $A_{11} \in [0, 1]^{Y_1 \times X_1}$  and  $A_{12} \in [0, 1]^{Y_1 \times X_2}$  come with their *approximation sequences*. They are increasing sequences of length  $\bar{\alpha} \leq \omega$ :

$$A_{11}^{(0)} \leq A_{11}^{(1)} \leq \dots \leq A_{11}^{(\bar{\alpha})} \in [0, 1]^{Y_1 \times X_1} \quad \text{and} \quad A_{12}^{(0)} \leq A_{12}^{(1)} \leq \dots \leq A_{12}^{(\bar{\alpha})} \in [0, 1]^{Y_1 \times X_2}$$

such that:

- (a) (**Approximate  $A_{11}$  and  $A_{12}$** ) We have  $A_{11}^{(\bar{\alpha})} = A_{11}$  and  $A_{12}^{(\bar{\alpha})} = A_{12}$ .
- (b) ( $A_{11}^{(\alpha)}$ ) For each  $\alpha \leq \bar{\alpha}$  and  $a \in \mathbf{A}$  we have:  $A_{11}^{(\alpha)} \cdot M_{\mathcal{X},1}(a) \leq M_{\mathcal{Y},1}(a) \cdot \begin{pmatrix} A_{11}^{(\alpha)} & A_{12}^{(\alpha)} \\ A_{21} & A_{22} \end{pmatrix}$ .
- (c) ( $A_{12}^{(\alpha)}$ , **base**) The 0-th approximant  $A_{12}^{(0)}$  is the zero matrix  $O$ .
- (d) ( $A_{12}^{(\alpha)}$ , **step**) For each  $\alpha < \bar{\alpha}$  and  $a \in \mathbf{A}$ :  $A_{12}^{(\alpha+1)} \cdot M_{\mathcal{X},2}(a) \leq M_{\mathcal{Y},1}(a) \cdot \begin{pmatrix} A_{11}^{(\alpha)} & A_{12}^{(\alpha)} \\ A_{21} & A_{22} \end{pmatrix}$ .
- (e) ( $A_{12}^{(\alpha)}$ , **limit**)  $(A_{12}^{(\omega)})_{y,x} = \sup_{\alpha' < \omega} (A_{12}^{(\alpha')})_{y,x}$  for each  $y \in Y_1$  and  $x \in X_2$ , in case  $\bar{\alpha} = \omega$ .

This notion is the combination of: 1) Kleisli simulation (see [UH17] and also Table 1(c) later) for mimicking one-step behaviors; and 2) progress measure [HSC16] that accounts for the nonlocal “fairness” constraint (Section 2). Indeed, Condition (2) and (3b) express the *invariant/gfp* intuition—note that (bi)simulation (without fairness) is a coinductive notion—while Condition (3c)–(3e) bears the *ranking function/lfp* flavor, mirroring the Cousot-Cousot approximation sequence  $\perp \sqsubseteq f(\perp) \sqsubseteq \dots$ .

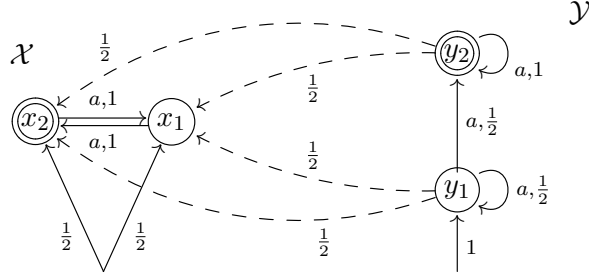
**Theorem 4.13** (soundness). *Assume  $\mathcal{Y}$  has a finite state space. Existence of a fair simulation (Definition 4.12) implies trace inclusion:  $L(\mathcal{X})(P) \leq L(\mathcal{Y})(P)$  for any measurable  $P \subseteq \mathbf{A}^{\omega}$ .*

The proof is presented later in Section 6, after we introduce coalgebraic machinery behind the definition of simulation.

We emphasize again that, differently from the nondeterministic setting, soundness of simulation is ensured only for *word* automata with a *finite* state space on the simulating side.

A (nontrivial) example of such a fair simulation is given below.

**Example 4.14.** Let  $\mathcal{X}$  and  $\mathcal{Y}$  be the PBWAs illustrated below.



We define  $A \in [0, 1]^{[y_1, y_2] \times \{x_1, x_2\}}$  by  $A_{y_i, x_j} = \frac{1}{2}$  for each  $i, j \in \{1, 2\}$ . Then  $A$  is a fair matrix simulation from  $\mathcal{X}$  to  $\mathcal{Y}$ . Here the approximation sequences  $A_{11}^{(0)} \sqsubseteq A_{11}^{(1)} \sqsubseteq \dots \sqsubseteq A_{11}^{(\omega)}$  and  $A_{12}^{(0)} \sqsubseteq A_{12}^{(1)} \sqsubseteq \dots \sqsubseteq A_{12}^{(\omega)}$  are given by  $A_{11}^{(i)} = (\frac{1}{2} - (\frac{1}{2})^{i+1}) \in [0, 1]^{[y_1] \times \{x_1\}}$  and  $A_{12}^{(i)} = (\frac{1}{2} - (\frac{1}{2})^{i+1}) \in [0, 1]^{[y_1] \times \{x_2\}}$  for each  $i$ .

## 5. COALGEBRAIC BACKGROUND

The fair simulation notions in Sections 3–4 (for nondeterminism and probability) may look different, but they arise from the same source, namely our coalgebraic study of Büchi automata [USH16].

**5.1. Modeling a System as a Function  $X \rightarrow TFX$ .** The conventional coalgebraic modeling of systems—as a function  $X \rightarrow FX$ —is known to capture *branching-time* semantics such as bisimilarity [Jac16, Rut00]. In contrast accepted languages of Büchi automata with nondeterministic or probabilistic branching constitute *linear-time* semantics; see [vG01] for the so-called *linear time-branching time spectrum*.

For the coalgebraic modeling of such linear-time semantics we follow the “Kleisli modeling” tradition [PT97, Jac04, HJS07]. Here a system is parametrized by a monad  $T$  and an endofunctor  $F$  on **Sets**: the former represents the *branching type* while the latter represents the (*linear-time*) *transition type*; and a system is modeled as a function of the type  $X \rightarrow TFX$ .<sup>5</sup>

A *monad*  $T$  is a construct from category theory [Mac98]: it is a functor  $T: \mathbb{C} \rightarrow \mathbb{C}$  equipped with *unit*  $\eta_X^T: X \rightarrow TX$  and *multiplication*  $\mu_X^T: T^2X \rightarrow TX$ , both given by arrows in  $\mathbb{C}$  for each object  $X \in \mathbb{C}$ , subject to some axioms. In this paper we use two examples  $T = \mathcal{P}, \mathcal{G}$ : the *powerset monad*  $\mathcal{P}$  (on the category **Sets** of sets and functions) for nondeterminism; and the *sub-Giry monad*  $\mathcal{G}$  (on **Meas** of measurable spaces and measurable functions) for probabilistic branching. The latter is a “sub” variant of the well-known *Giry monad* [Gir82].

<sup>5</sup> Another eminent approach to coalgebraic linear-time semantics is the *Eilenberg-Moore* one (see [JSS15, ABH<sup>+</sup>12] for example): notably in the latter a system is expressed as  $X \rightarrow FTX$ . The Eilenberg-Moore approach can be seen as a categorical generalization of *determinization* or the *powerset construction*. This however makes the approach hard to apply to *infinite* words or trees, since already for Büchi word automata, it is known that deterministic ones are less expressive than general, nondeterministic ones.

**Definition 5.1** (the monads  $\mathcal{P}$  and  $\mathcal{G}$ ). The *powerset monad*  $\mathcal{P}$  on **Sets** carries a set  $X$  to  $\mathcal{P}X = \{S \subseteq X\}$ , and a function  $f: X \rightarrow Y$  to  $\mathcal{P}f: \mathcal{P}X \rightarrow \mathcal{P}Y$ ,  $S \mapsto f[S] = \{f(x) \mid x \in S\}$ . For each set  $X$ , its unit  $\eta_X^{\mathcal{P}}: X \rightarrow \mathcal{P}X$  is given by the singleton map  $x \mapsto \{x\}$ ; and its multiplication  $\mu_X^{\mathcal{P}}: \mathcal{P}^2X \rightarrow \mathcal{P}X$  is given by union  $M \mapsto \bigcup_{A \in M} A$ .

The *sub-Giry monad*  $\mathcal{G}$  on **Meas** carries a measurable space  $(X, \mathfrak{F}_X)$  to  $(\mathcal{G}X, \mathfrak{F}_{\mathcal{G}X})$ , where  $\mathcal{G}X$  is the set of all *subprobability measures* on  $X$  and  $\mathfrak{F}_{\mathcal{G}X}$  is the smallest  $\sigma$ -algebra such that, for each  $S \in \mathfrak{F}_X$ , the function  $\text{ev}_S: \mathcal{G}X \rightarrow [0, 1]$  defined by  $\text{ev}_S(P) = P(S)$  is measurable. The action of  $\mathcal{G}$  on arrows is given by the pushforward measure: for  $f: X \rightarrow Y$ ,  $P \in \mathcal{G}X$  and  $T \in \mathfrak{F}_Y$ ,  $(\mathcal{G}f)(P)(T) = P(f^{-1}(T))$ . The unit  $\eta_X^{\mathcal{G}}: X \rightarrow \mathcal{G}X$  is given by the *Dirac measure*  $\eta_X^{\mathcal{G}}(x) = \delta_x$ ; and  $\mu_X^{\mathcal{G}}: \mathcal{G}^2X \rightarrow \mathcal{G}X$  is given by  $\Psi \mapsto (S \mapsto \int_{\mathcal{G}(X, \mathfrak{F}_X)} \text{ev}_S d\Psi)$ .

Intuitively  $\eta_X^T: X \rightarrow TX$  turns an element into a trivial branching while  $\mu_X^T: T^2X \rightarrow TX$  suppresses two successive branchings into one. See [HJS07] for further illustration.

For the other parameter  $F$ —for the type of linear-time behaviors—we use the following.

**Definition 5.2** (the functors  $F_{\Sigma}$  on **Sets** and  $F_A$  on **Meas**). Let  $\Sigma$  be a ranked alphabet. The functor  $F_{\Sigma}: \mathbf{Sets} \rightarrow \mathbf{Sets}$  carries a set  $X$  to  $F_{\Sigma}X = \coprod_{\sigma \in \Sigma} X^{|\sigma|}$ ; and a function  $f$  to  $\coprod_{\sigma \in \Sigma} f^{|\sigma|}$ . Let  $A$  be a countable alphabet, thought of as a measurable set with the discrete  $\sigma$ -algebra. The functor  $F_A = A \times (\_): \mathbf{Meas} \rightarrow \mathbf{Meas}$  carries a measurable space  $X$  to the product space  $A \times X$ ; and a measurable map  $f$  to  $\text{id}_A \times f$ .

Our system models in Sections 3–4 readily allow categorical modeling as arrows  $X \rightarrow TFX$ : the transition function of an NBTA (Definition 3.1) is a function  $X \rightarrow \mathcal{P}F_{\Sigma}X$ ; and the transition matrices of a PBWA (Definition 4.1) collectively give a (measurable) function  $X \rightarrow \mathcal{G}F_AX$ .

**5.2. Coalgebras in a Kleisli Category.** Given a monad  $T$  on a category  $\mathbb{C}$ , the standard construction of the *Kleisli category*  $\mathcal{Kl}(T)$  is defined as follows (see [Mac98] for example): its objects are those of  $\mathbb{C}$ ; its arrows  $f: X \rightarrow Y$  are precisely arrows  $f: X \rightarrow TY$  in  $\mathbb{C}$ ; and its identity and composition  $\odot$  are defined with the aid of unit  $\eta^T$  and multiplication  $\mu^T$ .<sup>6</sup> It is known that an arrow  $f: X \rightarrow Y$  in  $\mathbb{C}$  can be lifted to the Kleisli category  $\mathcal{Kl}(T)$  by the *Kleisli inclusion functor*  $J: \mathbf{Sets} \rightarrow \mathcal{Kl}(T)$  that is defined by  $f \mapsto \eta_Y \circ f$  [Mac98].

Intuitively a Kleisli arrow  $f: X \rightarrow Y$  is a function from  $X$  to  $Y$  *with  $T$ -branching*. Then a system dynamics  $X \rightarrow TFX$  with  $T$ -branching over linear-time  $F$ -behaviors is a Kleisli arrow  $X \rightarrow \overline{F}X$ , a (proper)  $\overline{F}$ -coalgebra in  $\mathcal{Kl}(T)$ . Here  $\overline{F}: \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  is a canonical lifting of  $F: \mathbb{C} \rightarrow \mathbb{C}$ , which is formally defined as follows.

**Definition 5.3.** For  $F: \mathbb{C} \rightarrow \mathbb{C}$ , a functor  $\overline{F}: \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  is called a *lifting* of a functor  $F: \mathbb{C} \rightarrow \mathbb{C}$  if  $\overline{F}X = FX$  and  $\overline{F} \circ J = J \circ F$ .

A canonical lifting can be explicitly described when  $T = \mathcal{P}$  and  $F = F_{\Sigma}$  on **Sets**, and when  $T = \mathcal{G}$  and  $F = F_A$  on **Meas**. See [HJS07, UH15] for example.

Studies of coalgebras  $X \rightarrow \overline{F}X$  are initiated in [PT97] and developed henceforth in [Jac04, HJS07, Cır10, KK13, UH17, UH15] for example, leading to the following *coalgebraic theory of trace and simulation*.

**(Table 1(a)).** In [HJS07] it is shown that, for  $T = \mathcal{P}$  (for nondeterminism) and  $\mathcal{D}$  (the *subdistribution* monad on **Sets** for discrete probabilities), and for a suitable functor  $F$  on

<sup>6</sup>For distinction we write  $\rightarrow$  for arrows in  $\mathcal{Kl}(T)$  (not  $\rightarrow$ ), and  $\odot$  for composition in  $\mathcal{Kl}(T)$  (not  $\circ$ ).

$$\begin{array}{ccc}
\begin{array}{c} \overline{FX} \xrightarrow{\overline{F}(\text{tr}(c))} \overline{FA} \\ c \uparrow = \uparrow J\alpha^{-1} \\ X \xrightarrow{\text{tr}(c)} A \end{array} & 
\begin{array}{c} \overline{FX} \xrightarrow{\overline{F}(\text{tr}^\infty(c))} \overline{FZ} \\ c \uparrow = \nu \uparrow J\zeta \\ X \xrightarrow{\text{tr}^\infty(c)} Z \end{array} & 
\begin{array}{c} \overline{FX} \xleftarrow{\overline{F}f} \overline{FY} \quad \overline{FX} \xrightarrow{\overline{F}b} \overline{FY} \\ c \uparrow \sqsubseteq_f \uparrow d \quad c \uparrow \sqsubseteq_b \uparrow d \\ X \xleftarrow{f} Y \quad X \xrightarrow{b} Y \end{array} \\
\text{(a) Coalgebraic finite trace:} & \text{(b) Coalgebraic infinitary trace:} & \text{(c) Coalgebraic fwd. and bwd. simulation:} \\
FA \xrightarrow{\alpha} A \text{ is an init. alg. in } \mathbf{Sets} & Z \xrightarrow{\zeta} FZ \text{ is a final coalg. in } \mathbf{Sets} & \text{here } c \text{ is simulating by } d
\end{array}$$

Table 1: Some known results in the coalgebraic theory of trace and simulation

**Sets**, an initial  $F$ -algebra  $\alpha: FA \cong A$  in **Sets** yields a final  $\overline{F}$ -coalgebra  $J\alpha^{-1}: A \rightarrow \overline{FA}$ . In case  $F = F_\Sigma$  an initial algebra is given by the set of all *finite*  $\Sigma$ -trees; and the unique morphism  $\text{tr}(c): X \rightarrow A$ —namely a function  $\text{tr}(c): X \rightarrow TA$ , see Table 1(a)—is nothing but the *finite trace semantics* of the automaton  $c: X \rightarrow \overline{FX}$ , capturing all the linear-time behaviors that *eventually terminate*.

**(Table 1(b))**. For *infinitary trace semantics* its coalgebraic characterization is more involved [Jac04, Cîr10]. Here we consider all possibly nonterminating linear-time behaviors of an automaton. In the above setting, and also for  $T = \mathcal{G}$  on **Meas**, it is shown that a final coalgebra  $\zeta: Z \cong FZ$ —we have  $Z \cong \text{Tree}_\Sigma$  when  $F = F_\Sigma$ —yields a *weakly final* coalgebra  $J\zeta$  in  $\mathcal{Kl}(T)$ . Given  $c$  there is thus at least one morphism from  $c$  to  $J\zeta$ ; there is also a *maximal* such  $\text{tr}^\infty(c)$ , and this is how we capture infinitary trace. In Table 1(b) we indicate this maximality by  $\nu$ .

**(Table 1(c))**. In [Has06] it is shown that *lax/oplax homomorphisms* (Table 1(c)) witness *finite trace inclusion*  $\text{tr}(c) \sqsubseteq \text{tr}(d)$ . When  $T = \mathcal{P}$  these notions specialize to *forward* and *backward simulation* in [LV95], namely binary relations that “mimic.” In [UH15] they are shown to witness *infinitary trace inclusion* too; this is the starting point of the current study of (forward) simulation for Büchi automata. Note that, when  $T = \mathcal{G}$ , our (forward) “simulation” is not a relation but a “function with probabilistic branching”  $f: Y \rightarrow \mathcal{G}X$ . The latter is roughly a matrix of dimension  $|Y| \times |X|$ ; and algorithms to find such are studied in [UH17].

**5.3. Coalgebraic Modeling of Büchi Automata.** In the above theory—and in the theory of coalgebra in general—the Büchi acceptance condition has long been considered a big challenge: its nonlocal character (“visit  $\odot$  infinitely often”) does not go along with the coalgebraic, local idea of behaviors that is centered around *homomorphisms* of coalgebras ( $f$  in the diagram).

$$\begin{array}{ccc}
FX & \xrightarrow{Ff} & FY \\
c \uparrow & & \uparrow d \\
X & \xrightarrow{f} & Y
\end{array}$$

Our answer [USH16] to the challenge, inspired by Table 1(b) and our recent [HSC16], consists of: 1) regarding the distinction of  $\bigcirc$  vs.  $\odot$  as a *partition*  $X = X_1 + X_2$  of the state space; and 2) introducing explicit  $\mu$ 's and  $\nu$ 's in commuting diagrams, hence regarding them as part of *equational systems* (Section 2). This forces our departure from the coalgebraic reasoning principle of *finality*—namely *existence* of a *unique* homomorphism—by moving from Table 1(a) to (5.3) below. We however believe this is a necessary step forward, for the theory of coalgebras to cope with its long-standing challenges like the Büchi condition and weak bisimilarity.

We review the part of the theory in [USH16] that is relevant to us.

**Definition 5.4** (Büchi  $(T, F)$ -system). Let  $T$  be a monad, and  $F$  be an endofunctor, both on some category  $\mathbb{C}$  with binary coproducts  $+$  and a nullary product  $1$ . Assume also that  $F$  lifts to  $\overline{F}: \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  (Definition 5.3).

A *Büchi  $(T, F)$ -system* is given by a tuple  $\mathcal{X} = ((X_1, X_2), c: X \rightarrow \overline{F}X, s: 1 \rightarrow X)$  where:

- $X_1$  and  $X_2$  are objects of  $\mathbb{C}$  (with the intuition that  $X_1 = \{\text{non-accepting states } \circlearrowleft\}$  and  $X_2 = \{\text{non-accepting states } \circlearrowright\}$ ), and we define  $X := X_1 + X_2$ ;
- $c: X \rightarrow \overline{F}X$  is an arrow in  $\mathcal{Kl}(T)$  for *dynamics*; and
- $s: 1 \rightarrow X$  is an arrow in  $\mathcal{Kl}(T)$  for *initial states*.

For each  $i \in \{1, 2\}$ , we define  $c_i: X_i \rightarrow \overline{F}X$  to be the restriction  $c \circ \kappa_i: X_i \rightarrow TFX$  of  $c$  along the coprojection  $\kappa_i: X_i \hookrightarrow X$ .

Thus a Büchi  $(T, F)$ -system is a (Kleisli) coalgebra  $X \rightarrow \overline{F}X$  augmented with the information on accepting and initial states. We can regard NBTAs and PBWAs as Büchi  $(T, F)$ -systems as follows; note that an arrow  $1 \rightarrow X$  in  $\mathcal{Kl}(\mathcal{G})$  is nothing but a probability subdistribution over  $X$ .

**Example 5.5.**

- (1) An NBTa  $\mathcal{X} = (X, \Sigma, \delta, I, \text{Acc})$  (Definition 3.1) gives rise to a Büchi  $(\mathcal{P}, F_\Sigma)$ -system  $\mathcal{X}' = ((X_1, X_2), c: X \rightarrow \overline{F}_\Sigma X, s: 1 \rightarrow X)$  that is defined by:
  - $X_1 = \text{Acc}$  and  $X_2 = X \setminus \text{Acc}$ ;
  - $c(x) = \delta(x)$ ; and
  - $s(*) = I$ .
- (2) A PBWA  $\mathcal{X} = (X, \mathbf{A}, M, \iota, \text{Acc})$  (Definition 4.1) gives rise to a Büchi  $(\mathcal{G}, F_{\mathbf{A}})$ -system  $\mathcal{X}' = ((X_1, X_2), c: X \rightarrow \overline{F}_{\mathbf{A}} X, s: 1 \rightarrow X)$  that is defined by:
  - $X_1 = (\text{Acc}, \mathcal{P}\text{Acc})$  and  $X_2 = (X \setminus \text{Acc}, \mathcal{P}(X \setminus \text{Acc}))$ ;
  - $c(x)(\{(a, x')\}) = (M(a))_{x, x'}$ ; and
  - $s(*)({x}) = \iota_x$ .

Here  $c$  and  $s$  are well-defined as  $X_1 + X_2 \in \mathbf{Meas}$  is equipped with the discrete  $\sigma$ -algebra.

The next is the main theorem of [USH16].<sup>7</sup> Recall that  $\text{Tree}_\Sigma$  is the set of (possibly infinite)  $\Sigma$ -trees (Section 3.1); it carries a final coalgebra  $\zeta: \text{Tree}_\Sigma \xrightarrow{\cong} F_\Sigma(\text{Tree}_\Sigma)$  in **Sets**. We will be using natural orders  $\sqsubseteq_{X, Y}$  on the homsets  $\mathcal{Kl}(\mathcal{P})(X, Y)$  and  $\mathcal{Kl}(\mathcal{G})(X, Y)$ , given by inclusion and pointwise extension of the order on  $[0, 1]$ , respectively. Namely,

$$\begin{aligned} f \sqsubseteq_{X, Y} g &\stackrel{\text{def}}{\iff} \forall x \in X. f(x) \subseteq g(x) && \text{for } T = \mathcal{P} \quad \text{and} \\ f \sqsubseteq_{X, Y} g &\stackrel{\text{def}}{\iff} \forall x \in X. \forall A \in \mathfrak{F}_Y. f(x)(A) \leq g(x)(A) && \text{for } T = \mathcal{G}. \end{aligned} \tag{5.1}$$

**Theorem 5.6** [USH16].

- (1) Let  $\mathcal{X} = ((X_1, X_2), c, s)$  be a Büchi  $(\mathcal{P}, F_\Sigma)$ -system. Consider an equational system

$$u_1 =_\mu (J\zeta)^{-1} \odot \overline{F}_\Sigma[u_1, u_2] \odot c_1, \quad u_2 =_\nu (J\zeta)^{-1} \odot \overline{F}_\Sigma[u_1, u_2] \odot c_2 \tag{5.2}$$

<sup>7</sup>In fact this is a special case of the main theorem because the original theorem considers *parity*  $(\mathcal{P}, F_\Sigma)$ -systems, which generalizes Büchi  $(\mathcal{P}, F_\Sigma)$ -systems and is identified with parity tree automata. Note that the Büchi acceptance condition is a special case of the parity acceptance condition.

where  $u_i$  ranges over the homset  $\mathcal{Kl}(\mathcal{P})(X_i, \text{Tree}_\Sigma)$  for  $i \in \{1, 2\}$ . Diagrammatically:

$$\begin{array}{ccc}
F_\Sigma X \xrightarrow{\overline{F_\Sigma}[u_1, u_2]} F_\Sigma(\text{Tree}_\Sigma) & F_\Sigma X \xrightarrow{\overline{F_\Sigma}[u_1, u_2]} F_\Sigma(\text{Tree}_\Sigma) & \\
c_1 \uparrow \quad \quad \quad =_\mu \quad \quad \quad \cong \uparrow J\zeta & c_2 \uparrow \quad \quad \quad =_\nu \quad \quad \quad \cong \uparrow J\zeta & (5.3) \\
X_1 \xrightarrow{\quad \quad \quad u_1 \quad \quad \quad} \text{Tree}_\Sigma & X_2 \xrightarrow{\quad \quad \quad u_2 \quad \quad \quad} \text{Tree}_\Sigma & .
\end{array}$$

- (a) The equational system has a solution, denoted by  $\text{tr}^B(c_i): X_i \rightarrow \text{Tree}_\Sigma$  for  $i \in \{1, 2\}$ .
- (b) Let  $\text{tr}^B(\mathcal{X}) := (\{*\} = 1 \xrightarrow{s} X = X_1 + X_2 \xrightarrow{[\text{tr}^B(c_1), \text{tr}^B(c_2)]} \text{Tree}_\Sigma)$  be a composite in  $\mathcal{Kl}(\mathcal{P})$ . In case  $\mathcal{X}$  is induced by an NBTA, the set  $\text{tr}^B(\mathcal{X})(*) \subseteq \text{Tree}_\Sigma$  coincides with the (Büchi) language  $L(\mathcal{X})$  of  $\mathcal{X}$  (Definition 3.8).
- (2) Let  $\mathcal{X}$  be a Büchi  $(\mathcal{G}, F_A)$ -system, and consider the same equational system as (5.2), but with  $\mathcal{G}, F_A, A^\omega$  replacing  $\mathcal{P}, F_\Sigma, \text{Tree}_\Sigma$ . Then:
- (a) The equational system has a solution.
- (b) Let  $\mathcal{X}$  be induced by a PBWA (Definition 4.1). For the same composite  $\text{tr}^B(\mathcal{X}): 1 \rightarrow A^\omega$  as above we have  $\text{tr}^B(\mathcal{X})(*) = L(\mathcal{X}) \in \mathcal{G}(A^\omega)$ , the Büchi language of the PBWA (Section 4).  $\square$

In the proof of the above theorem, (1a) and (2a) are proved using Proposition 2.3. More concretely, if  $T = \mathcal{P}$  and  $F = F_\Sigma$  then Condition (a) of Proposition 2.3 is satisfied by the equational system. In contrast, if  $T = \mathcal{G}$  and  $F = F_A$  then Condition (b) is satisfied.

## 6. COALGEBRAIC ACCOUNT ON FAIR SIMULATIONS AND SOUNDNESS PROOFS

Here we lay out our coalgebraic study of fair simulations. We will be firstly led to a simulation notion “with dividing” that is coalgebraically neat but not desirable from a practical viewpoint. Circumventing the dividing construct we obtain the simulation notions that we have presented in Sections 3–4.

The last part of circumventing dividing is different for  $T = \mathcal{P}$  and  $\mathcal{G}$ ; this is why we have different definitions of simulation. While one would hope for uniformity, we suspect it to be hard, for the following reason. We observed [UH15] that the characterization of infinite trace (Table 1(b)) is true for  $T = \mathcal{P}$  and  $\mathcal{G}$ , but because of different categorical machineries. Since infinite trace is a special case of Büchi acceptance (where every state is accepting) and our soundness proof should rely on its characterization, we expect that this sharp contrast would still stand.

### 6.1. Cppo-enriched Categories and Functors; Codomain Restrictions and Joins.

In this section, we review four categorical constructs that are used in the definition of our categorical simulation notion.

**6.1.1. Cppo-enriched category and Cppo-enriched functor.** Recall that in the categorical definition of Büchi languages, we used a partial order  $\sqsubseteq_{X,Y}$  on each homset  $\mathcal{Kl}(T)(X, Y)$ . The first two notions—**Cppo-enriched category** and **Cppo-enriched functor**, see e.g. [Bor94]—add certain assumptions to the ordered structure. The same assumptions are also used in [HJS07] where finite trace semantics of nondeterministic and probabilistic systems are captured categorically.

**Definition 6.1** (**Cppo**-enriched category and **Cppo**-enriched functor). A category  $\mathbb{C}$  is called a **Cppo**-enriched category if it satisfies the following conditions:

- (1) Each homset  $\mathbb{C}(X, Y)$  carries a partial order  $\sqsubseteq_{X, Y}$ . Moreover each homset  $\mathbb{C}(X, Y)$  is a *pointed cpo* with respect to the order, i.e. it has the least element  $\perp_{X, Y}$  and each increasing sequence  $f_0 \sqsubseteq_{X, Y} f_1 \sqsubseteq_{X, Y} \dots \in \mathbb{C}(X, Y)$  has the least upper bound  $\bigsqcup_{i \in \omega} f_i : X \rightarrow Y$ .
- (2) For each  $X, Y, Z \in \mathbb{C}$ , the composition  $(\_ \circ \_) : \mathbb{C}(Y, Z) \times \mathbb{C}(X, Y) \rightarrow \mathbb{C}(X, Z)$  is monotone with respect to the product order.
- (3) The composition  $\circ$  is  $\omega$ -continuous. That is, for an increasing sequence  $f_0 \sqsubseteq_{X, Y} f_1 \sqsubseteq_{X, Y} \dots$  of arrows,

$$\left(\bigsqcup_{i < \omega} f_i\right) \circ g = \bigsqcup_{i < \omega} (f_i \circ g) \quad \text{and} \quad h \circ \left(\bigsqcup_{i < \omega} f_i\right) = \bigsqcup_{i < \omega} (h \circ f_i). \quad (6.1)$$

Let  $\mathbb{C}$  be a **Cppo**-enriched category. A functor  $F : \mathbb{C} \rightarrow \mathbb{C}$  is called a **Cppo**-enriched functor if it satisfies the following conditions.

- (a) It is *locally monotone*, that is, for each  $X, Y \in \mathbb{C}$  and  $f, g : X \rightarrow Y$ ,  $f \sqsubseteq_{X, Y} g$  implies  $Ff \sqsubseteq_{FX, FY} Fg$ .
- (b) It is *locally  $\omega$ -continuous*, that is, for each  $X, Y \in \mathbb{C}$  and increasing sequence  $f_0 \sqsubseteq_{X, Y} f_1 \sqsubseteq_{X, Y} \dots \in \mathbb{C}(X, Y)$ , we have  $F(\bigsqcup_{i < \omega} f_i) = \bigsqcup_{i < \omega} (Ff_i)$ .

If confusion is unlikely, we omit subscripts and just write  $\sqsubseteq$  and  $\perp$  for  $\sqsubseteq_{X, Y}$  and  $\perp_{X, Y}$ .

**Remark 6.2.** In a definition of **Cppo**-enriched category, monotonicity of compositions (Condition (2) in the definition above) is often omitted (see [HJS07, BMSZ14] for example). However we require it explicitly to ensure that if  $(f_i)_{i \in \omega}$  is an increasing sequence then  $(f_i \circ g)_{i \in \omega}$  and  $(h \circ f_i)_{i \in \omega}$  are also increasing sequences and hence the suprema  $\bigsqcup_{i < \omega} (f_i \circ g)$  and  $\bigsqcup_{i < \omega} (h \circ f_i)$  in (6.1) in the definition is well-defined. We require a **Cppo**-enriched functor to be locally monotone (Condition (a) in the definition) for the same reason.

**Remark 6.3.** The notions of **Cppo**-enriched category and **Cppo**-enriched functor are instances of well-known categorical notions of  $\mathbb{V}$ -enriched category and  $\mathbb{V}$ -enriched functor (see [Bor94] for example). Later in the soundness proof of our categorical fair simulation, we will be assuming  $\mathcal{Kl}(T)$  and  $\overline{F}$  to be **Cppo**-enriched. We do so for conceptual simplicity: technically speaking this is stronger than needed, since  $\omega$ -continuity of composition is not used in the proof.

It is not so hard to see that  $\mathcal{Kl}(\mathcal{P})$  and  $\mathcal{Kl}(\mathcal{G})$  are both **Cppo**-enriched categories and  $\overline{F}_\Sigma$  and  $\overline{F}_A$  are both **Cppo**-enriched functors, with respect to the orders in (5.1).

**Remark 6.4.** Let  $T$  be a monad and  $F$  be a functor with a final coalgebra  $\zeta : Z \rightarrow FZ$ . If  $\mathcal{Kl}(T)$  and  $\overline{F}$  are **Cppo**-enriched then the equational system (5.3) (with  $T$ ,  $F$  and  $Z$  replacing  $\mathcal{P}$ ,  $F_\Sigma$  and  $\text{Tree}_\Sigma$ ) satisfies the assumptions (i) and (ii) in Theorem 2.7. Therefore by Theorem 2.7, progress measures for the equational system satisfy soundness and completeness.

There is another way to ensure soundness and completeness of progress measures for the equational system. In the original version of Theorem 2.7 in [HSC16], instead of the assumptions (i) and (ii), the following assumption is required:

- for each  $i \in [1, m]$ , the poset  $L_i$  is a complete lattice.

This implies that soundness and completeness of progress measures for the equational system (5.3) are satisfied if we have the following condition:

(‡) each homset of  $\mathcal{Kl}(T)$  carries a complete lattice; Kleisli compositions are monotone; and  $\overline{F}$  is locally monotone.

However, as we have mentioned in Remark 2.6, a homset of  $\mathcal{Kl}(\mathcal{G})$  does not necessarily carry a complete lattice, and hence  $T = \mathcal{G}$  does not satisfy (‡) above.

**Example 6.5.** A homset of  $\mathcal{Kl}(\mathcal{G})$  is not necessarily a dcpo; here is a counterexample.

Let  $\mathfrak{F}_{[0,1]}$  be the  $\sigma$ -algebra over the unit interval  $[0, 1]$  consisting of the Borel sets (see [ADD00] for example). It is known that there exists  $V \subseteq [0, 1]$  such that  $V \notin \mathfrak{F}_{[0,1]}$  (see [Her06] for example). It is easy to see that  $\mathcal{G}(1, \mathcal{P}1) \cong ([0, 1], \mathfrak{F}_{[0,1]})$ .

We define a set  $\mathfrak{A}_V \subseteq \mathcal{Kl}(\mathcal{G})(([0, 1], \mathfrak{F}_{[0,1]}), (1, \mathcal{P}1))$  of Kleisli arrows by

$$\mathfrak{A}_V := \{ \chi_X : ([0, 1], \mathfrak{F}_{[0,1]}) \rightarrow \mathcal{G}(1, \mathcal{P}1) \mid X \in \mathfrak{F}_{[0,1]} \text{ and } X \subseteq V \}.$$

Here  $\chi_X$  denotes the characteristic function of  $X$ , that is,  $\chi_X(x) = 1$  if  $x \in X$  and  $\chi_X(x) = 0$  otherwise. Note that a Kleisli arrow  $\chi_X : ([0, 1], \mathfrak{F}_{[0,1]}) \rightarrow \mathcal{G}(1, \mathcal{P}1)$  is a measurable function  $\chi_X : ([0, 1], \mathfrak{F}_{[0,1]}) \rightarrow ([0, 1], \mathfrak{F}_{[0,1]})$ . It is easy to see that  $\chi_X \sqsubseteq \chi_{X'}$  if and only if  $X \subseteq X'$ .

Assume that  $\mathcal{Kl}(\mathcal{G})(([0, 1], \mathfrak{F}_{[0,1]}), (1, \mathcal{P}1))$  is a dcpo. Then there exists the least upper bound  $\bigsqcup_{X \subseteq V} \chi_X : ([0, 1], \mathfrak{F}_{[0,1]}) \rightarrow \mathcal{G}(1, \mathcal{P}1)$  of  $\mathfrak{A}_V$ .

Let  $V' := (\bigsqcup_{X \subseteq V} \chi_X)^{-1}(\{1\})$ . Then as  $\{1\} \in \mathfrak{F}_{[0,1]}$  and  $\bigsqcup_{X \subseteq V} \chi_X$  is a measurable function, we have  $V' \in \mathfrak{F}_{[0,1]}$ . Moreover as  $\bigsqcup_{X \subseteq V} \chi_X$  is an upper bound of  $\mathfrak{A}_V$ , we have  $V \subseteq V'$ . Therefore by  $V \notin \mathfrak{F}_{[0,1]}$ , there exists  $v \in V'$  such that  $V \subseteq V' \setminus \{v\}$ . As  $\{v\}, V' \in \mathfrak{F}_{[0,1]}$ , we have  $V' \setminus \{v\} \in \mathfrak{F}_{[0,1]}$ . It is easy to see that  $\chi_{V' \setminus \{v\}}$  is an upper bound of  $\mathfrak{A}_V$ . This contradicts the fact  $\bigsqcup_{X \subseteq V} \chi_X$  is the least upper bound of  $\mathfrak{A}_V$ . Hence  $\mathcal{Kl}(\mathcal{G})(([0, 1], \mathfrak{F}_{[0,1]}), (1, \mathcal{P}1))$  is not a dcpo.

**6.1.2. Codomain Restriction and Codomain Join.** The other two notions we describe in Section 6.1 are codomain restriction and codomain join of Kleisli arrows in  $\mathcal{Kl}(T)$ . The latter combines two arrows  $f_1 : X \rightarrow Y_1$  and  $f_2 : X \rightarrow Y_2$  into  $f : X \rightarrow Y_1 + Y_2$  while the former does the inverse. We note that the former operation is not necessarily total (see Example 6.8). It is known that if the monad  $T$  satisfies a certain condition, then its Kleisli category comes with the two operations.

**Definition 6.6** [Cîr13, Jac10]. Let  $\mathbb{C}$  be a category with an initial object  $0$ , binary products and binary coproducts. A monad  $T$  on  $\mathbb{C}$  is called a *partially additive monad* if it satisfies the following conditions:

- (1) The object  $T0$  is a final object in  $\mathbb{C}$ .<sup>8</sup>
- (2) Let  $X_1, X_2 \in \mathbb{C}$ . We define  $p_1 : X_1 + X_2 \rightarrow TX_1$  and  $p_2 : X_1 + X_2 \rightarrow TX_2$  by  $p_1 := [\eta_{X_1}, \perp_{X_2, X_1}]$  and  $p_2 := [\perp_{X_1, X_2}, \eta_{X_2}]$ . Here for each  $X, Y \in \mathbb{C}$ ,  $\perp_{X, Y} : X \rightarrow TY$  is given by  $\perp_{X, Y} := X \xrightarrow{!_X} T0 \xrightarrow{T!_Y} TY$  where  $!_X : X \rightarrow T0$  and  $!_T : 0 \rightarrow Y$  denote the unique arrows (see also Remark 6.9).

We require the following arrow be a monomorphism.

$$T(X_1 + X_2) \xrightarrow{\langle \mu_{X_1} \circ T p_1, \mu_{X_2} \circ T p_2 \rangle} TX_1 \times TX_2$$

<sup>8</sup>This implies that  $0$  is both an initial and final object in  $\mathcal{Kl}(T)$ . Such  $0$  is called a *zero object*.



**Definition 6.7** (codomain restriction and codomain join, [Cîr13, Jac10]). Let  $T$  be a partially additive monad. Then  $\mathcal{Kl}(T)$  comes with two operations on arrows called *codomain restriction* and *codomain join*. Codomain restriction takes an arrow  $g: V \rightarrow X_1 + X_2$  and returns  $g \upharpoonright^{X_i}: V \rightarrow X_i$  for  $i \in \{1, 2\}$ . Here  $g \upharpoonright^{X_i}$  is defined by

$$g \upharpoonright^{X_i}: V \xrightarrow{g} T(X_1 + X_2) \xrightarrow{\langle \mu_{X_1} \circ Tp_1, \mu_{X_2} \circ Tp_2 \rangle} TX_1 \times TX_2 \xrightarrow{\pi_i} TX_i.$$

Codomain join is a *partial* operation that takes a pair  $g_1: V \rightarrow X_1$  and  $g_2: V \rightarrow X_2$ , and returns a (necessarily unique) arrow  $\langle\langle g_1, g_2 \rangle\rangle: V \rightarrow X_1 + X_2$  such that

$$\langle g_1, g_2 \rangle = \langle \mu_{X_1} \circ Tp_1, \mu_{X_2} \circ Tp_2 \rangle \circ \langle\langle g_1, g_2 \rangle\rangle.$$

The situation is illustrated below.

$$\begin{array}{ccc} T(X_1 + X_2) & \xrightarrow{\langle \mu_{X_1} \circ Tp_1, \mu_{X_2} \circ Tp_2 \rangle} & TX_1 \times TX_2 \\ & \searrow \langle\langle g_1, g_2 \rangle\rangle & \uparrow \langle g_1, g_2 \rangle \\ & & V \end{array}$$

These operations may look unfamiliar, but  $T = \mathcal{P}, \mathcal{G}$  are partially additive monads, and codomain joins and operations are given by suitably restricting/joining subsets/distributions.

**Example 6.8.** For  $T = \mathcal{P}$ , we can define codomain restrictions and codomain joins by

$$g \upharpoonright^{X_i}(v) = \{x \in X_i \mid x \in g(v)\} \quad \text{and} \quad \langle\langle g_1, g_2 \rangle\rangle(v) = g_1(v) \cup g_2(v).$$

For  $T = \mathcal{G}$ , the definitions are as follows.

$$g \upharpoonright^{X_i}(v)(A) = g(v)(A), \quad \text{and}$$

$$\langle\langle g_1, g_2 \rangle\rangle(v)(A) = \begin{cases} g_1(v)(A \cap X_1) + g_2(v)(A \cap X_2) & (g_1(v)(A \cap X_1) + g_2(v)(A \cap X_2) \leq 1) \\ \text{undefined} & (\text{otherwise}). \end{cases}$$

Note that codomain join for  $T = \mathcal{G}$  is a partial operation in the sense that it is not always defined.

**Remark 6.9.** Let  $T$  be a partially additive monad such that  $\mathcal{Kl}(T)$  is a **Cppo**-enriched category. We have used the same symbol  $\perp_{X,Y}$  for (1) the least element in a homset of a **Cppo**-enriched category (Definition 6.1) and (2) the arrow  $T \text{!}_{Y \circ !}_X$  in the Kleisli category of a partially additive monad. In the next section we assume that composition in  $\mathcal{Kl}(T)$  is left-strict, i.e.  $\perp_{X,Y} \odot g = \perp_{Z,Y}$  for each  $g: Z \rightarrow Y$ , where  $\perp_{X,Y}$  and  $\perp_{Z,Y}$  refer to the least arrows (see Theorem 6.13). It is easy to see that under this assumption, the arrow  $T \text{!}_{Y \circ !}_X$  coincides with the least element in  $\mathcal{Kl}(T)(X, Y)$ . This justifies overriding the symbol  $\perp$ .

We conclude this section with some properties of codomain restrictions and joins. They are proved by easy diagram-chasing.

**Lemma 6.10.** *Let  $T$  be a partially additive monad. We have the following:*

- (1) *Codomain restrictions and codomain joins are partially mutually inverse, in the following sense. Given  $g: V \rightarrow X_1 + X_2$ , the codomain join  $\langle\langle g \upharpoonright^{X_1}, g \upharpoonright^{X_2} \rangle\rangle$  is always defined and equal to  $g$ . Conversely, provided that  $\langle\langle g_1, g_2 \rangle\rangle$  is defined, we have  $(\langle\langle g_1, g_2 \rangle\rangle) \upharpoonright^{X_i} = g_i$  for  $i \in \{1, 2\}$ .*

- (2) For  $f : W \rightarrow V$ ,  $g_1 : V \rightarrow X_1$ ,  $g_2 : V \rightarrow X_2$ ,  $h_1 : X_1 \rightarrow Y_1$  and  $h_2 : X_2 \rightarrow Y_2$  such that  $\langle\langle g_1, g_2 \rangle\rangle$  is defined, we have:

$$\langle\langle g_1, g_2 \rangle\rangle \odot f = \langle\langle g_1 \circ f, g_2 \circ f \rangle\rangle \quad \text{and} \quad (h_1 + h_2) \circ \langle\langle g_1, g_2 \rangle\rangle = \langle\langle h_1 \circ g_1, h_2 \circ g_2 \rangle\rangle.$$

- (3) For  $g : V \rightarrow X$ ,  $\langle\langle g, \perp_{V,X} \rangle\rangle$  and  $\langle\langle \perp_{V,X}, g \rangle\rangle$  are always defined and we have

$$[\text{id}_X, \text{id}_X] \odot \langle\langle g, \perp_{V,X} \rangle\rangle = [\text{id}_X, \text{id}_X] \odot \langle\langle \perp_{V,X}, g \rangle\rangle = g. \quad \square$$

**6.2. Coalgebraic Fair Simulation with Dividing.** We make the following requirements in this section so that our definitions will make sense.

**Assumption 6.11.** *In this section we assume the following conditions on  $T$  and  $F$  on  $\mathbb{C}$ .*

- (1) The functor  $F$  has a final coalgebra  $\zeta : Z \xrightarrow{\cong} FZ$  in  $\mathbb{C}$ .
- (2) The functor  $F : \mathbb{C} \rightarrow \mathbb{C}$  lifts to  $\bar{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  (see Definition 5.3).
- (3) The Kleisli category  $\mathcal{Kl}(T)$  and the lifting  $\bar{F} : \mathcal{Kl}(T) \rightarrow \mathcal{Kl}(T)$  of  $F$  are both **Cppo**-enriched (Definition 6.1).
- (4) The monad  $T$  is a partially additive monad. Moreover the codomain join is downward closed. That is, for  $f_i, g_i : V \rightarrow X_i$  such that  $f_i \sqsubseteq g_i$  for each  $i \in I$ , if  $\langle\langle g_i \rangle\rangle_{i \in I}$  is defined, then so is  $\langle\langle f_i \rangle\rangle_{i \in I}$ .
- (5) Codomain restriction  $(\_) \upharpoonright^{X_i}$ , codomain join  $\langle\langle \_, \_ \rangle\rangle$  and cotupling  $[\_, \_]$  of Kleisli arrows are all monotone with respect to the order  $\sqsubseteq$ .

Note that by Definition 6.7, Condition (4) implies that  $\mathcal{Kl}(T)$  comes with codomain restrictions and joins.

With the help of codomain restrictions/joins we define a categorical fair simulation.

**Definition 6.12** ((forward) fair simulation with dividing). Let  $T$  and  $F$  be subject to Assumption 6.11;  $\mathcal{X} = ((X_1, X_2), c, s)$  and  $\mathcal{Y} = ((Y_1, Y_2), d, t)$  be Büchi  $(T, F)$ -systems; and  $\bar{\alpha}$  be an ordinal. A *(forward,  $\bar{\alpha}$ -bounded) fair simulation with dividing* from  $\mathcal{X}$  to  $\mathcal{Y}$  is an arrow  $f : Y \rightarrow X$  in  $\mathcal{Kl}(T)$  subject to the following conditions. Below, for simplicity, a domain-and-codomain restriction  $(f \odot \kappa_j) \upharpoonright^{X_i} : Y_j \rightarrow X_i$  (Definition 6.7) shall be denoted by  $f_{ji}$ ; and we refer to  $f_{11}, f_{12}, f_{21}, f_{22}$  as *components* of a fair simulation  $f$ .

- (1) The arrow  $f : Y \rightarrow X$  is a forward simulation from  $\mathcal{X}$  to  $\mathcal{Y}$  in the sense of [Has06] (see also Table 1(c)). That is:  $c \odot f \sqsubseteq \bar{F}f \odot d$  and  $s \sqsubseteq f \odot t$ .
- (2) The components  $f_{11} : Y_1 \rightarrow X_1$  and  $f_{12} : Y_1 \rightarrow X_2$  come with a *dividing*  $d_{11}, d_{12}$  of the component  $d_1 : Y_1 \rightarrow \bar{F}Y$  of  $d$ , and *approximation sequences*. The former is a pair  $d_{11}, d_{12} : Y_1 \rightarrow \bar{F}Y$  such that  $[\text{id}_{\bar{F}Y}, \text{id}_{\bar{F}Y}] \odot \langle\langle d_{11}, d_{12} \rangle\rangle = d_1$ . The latter are (possibly transfinite) increasing sequences of length  $\bar{\alpha}$ :

$$f_{11}^{(0)} \sqsubseteq f_{11}^{(1)} \sqsubseteq \cdots \sqsubseteq f_{11}^{(\bar{\alpha})} : Y_1 \rightarrow X_1, \quad \text{and} \quad f_{12}^{(0)} \sqsubseteq f_{12}^{(1)} \sqsubseteq \cdots \sqsubseteq f_{12}^{(\bar{\alpha})} : Y_1 \rightarrow X_2, \quad \text{such that}$$

- (a) **(Approximate  $f_{11}$  and  $f_{12}$ )** We have  $f_{11}^{(\bar{\alpha})} = f_{11}$  and  $f_{12}^{(\bar{\alpha})} = f_{12}$ .
- (b)  $(f_{11}^{(\alpha)})$  For each ordinal  $\alpha$  such that  $\alpha \leq \bar{\alpha}$ , the inequality (6.2) below holds. Note that the required codomain joins do exist.
- (c)  $(f_{12}^{(\alpha)}, \text{the base case})$  For the 0-th approximant, we have  $f_{12}^{(0)} = \perp$ .
- (d)  $(f_{12}^{(\alpha)}, \text{the step case})$  For each ordinal  $\alpha$  such that  $\alpha < \bar{\alpha}$ , the inequality (6.3) holds.

- (e) ( $f_{12}^{(\alpha)}$ , **the limit case**) If  $\alpha$  is a limit ordinal, then the supremum  $\bigsqcup_{\alpha' < \alpha} f_{12}^{(\alpha')}$  exists and  $f_{12}^{(\alpha)} \sqsubseteq \bigsqcup_{\alpha' < \alpha} f_{12}^{(\alpha')}$ .

$$\begin{array}{ccc} FY & \xrightarrow{\overline{F}[\langle\langle f_{11}^{(\alpha)}, f_{12}^{(\alpha)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\ d_{11} \uparrow & \sqsupseteq & \uparrow c_1 \\ Y_1 & \xrightarrow{f_{11}^{(\alpha)}} & X_1 \end{array} \quad (6.2)$$

$$\begin{array}{ccc} FY & \xrightarrow{\overline{F}[\langle\langle f_{11}^{(\alpha)}, f_{12}^{(\alpha)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\ d_{12} \uparrow & \sqsupseteq & \uparrow c_2 \\ Y_1 & \xrightarrow{f_{12}^{(\alpha+1)}} & X_2 \end{array} \quad (6.3)$$

In the definition above, note the direction: a simulation from  $\mathcal{X}$  to  $\mathcal{Y}$  has type  $Y \rightarrow X$ .

**Theorem 6.13** (soundness). *Let  $\bar{\alpha}$  be an ordinal. Assume Assumption 6.11 and the following.*

- (6) *For an arbitrary Büchi  $(T, F)$ -system  $\mathcal{X}$ , the equational system (5.2), with  $F, Z$  replacing  $F_\Sigma, \text{Tree}_\Sigma$ , has a (necessarily unique) solution  $\text{tr}^B(c_1), \text{tr}^B(c_2)$ .*
- (7) *The Kleisli composition  $\odot$  is both left and right-strict:  $\perp \odot f = \perp$  and  $f \odot \perp = \perp$ .*
- (8) *For each limit ordinal  $\alpha \leq \bar{\alpha}$ , post-composition in  $\mathcal{Kl}(T)$  is  $\alpha$ -continuous, i.e. if the supremum  $\bigsqcup_{i < \alpha} f_i$  exists then  $\bigsqcup_{i < \omega} (g \odot f_i)$  also exists and  $g \odot (\bigsqcup_{i < \alpha} f_i) = \bigsqcup_{i < \omega} (g \odot f_i)$ .*

*Then a fair  $\bar{\alpha}$ -bounded simulation with dividing, from one Büchi  $(T, F)$ -system  $\mathcal{X}$  to another  $\mathcal{Y}$ , witnesses trace inclusion  $\text{tr}^B(\mathcal{X}) \sqsubseteq \text{tr}^B(\mathcal{Y}) : 1 \rightarrow TZ$ .*

This theorem follows immediate from the following lemma.

**Lemma 6.14.** *Assume Assumption 6.11 and the assumptions (6)–(8) in Thm. 6.13. Let  $f : Y \rightarrow X$  be a forward,  $\bar{\alpha}$ -bounded simulation with dividing from  $\mathcal{X} = ((X_1, X_2), c, s)$  and  $\mathcal{Y} = ((Y_1, Y_2), d, t)$  (Definition 6.12). We define arrows  $\text{tr}^B(c_1) : X_1 \rightarrow Z$ ,  $\text{tr}^B(c_2) : X_2 \rightarrow Z$ ,  $\text{tr}^B(d_1) : Y_1 \rightarrow Z$  and  $\text{tr}^B(d_2) : Y_2 \rightarrow Z$  as in Theorem 5.6. Then we have:*

$$[\text{tr}^B(c_1), \text{tr}^B(c_2)] \odot \langle\langle f_{11}, f_{12} \rangle\rangle \sqsubseteq \text{tr}^B(d_1), \quad \text{and} \quad [\text{tr}^B(c_1), \text{tr}^B(c_2)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle \sqsubseteq \text{tr}^B(d_2).$$

To prove this lemma we need two sublemmas.

**Sublemma 6.15.** *We assume that  $T$  and  $F$  satisfy Assumption 6.11 and the assumptions in Theorem 6.13. We further assume the situation in Definition 6.12. Let  $d_{11}, d_{12} : X_1 \rightarrow \overline{F}X$  be the dividing of  $d_1 : X_1 \rightarrow \overline{F}X$ . Recall that  $\text{tr}^B(c_1) : X_1 \rightarrow Z$  and  $\text{tr}^B(c_2) : X_2 \rightarrow Z$  are given by the solutions  $u_1^{\text{sol}}$  and  $u_2^{\text{sol}}$  of the following equational system. (Theorem 5.6).*

$$\begin{aligned} u_1 &=_{\mu} (J\zeta)^{-1} \odot \overline{F}[u_1, u_2] \odot c_1 \in \mathcal{Kl}(T)(X_1, Z) \\ u_2 &=_{\nu} (J\zeta)^{-1} \odot \overline{F}[u_1, u_2] \odot c_2 \in \mathcal{Kl}(T)(X_2, Z) \end{aligned} \quad (6.4)$$

*By completeness of progress measure (Theorem 2.7.2), there exists a progress measure*

$$p_{\mathcal{X}} = ((\overline{\beta}_1), (u_1(\beta_1) : X_1 \rightarrow Z, u_2(\beta_1) : X_2 \rightarrow Z))_{\beta_1 \leq \overline{\beta}_1}$$

*for (6.4) such that  $\overline{\beta}_1 \leq \omega$ ,  $u_1(\overline{\beta}_1) = \text{tr}^B(c_1)$  and  $u_2 = \text{tr}^B(c_2)$ . We define two ordinals  $\overline{\gamma}_1$  and  $\overline{\gamma}_2$  by  $\overline{\gamma}_1 = \overline{\beta}_1$  and  $\overline{\gamma}_2 = \bar{\alpha}$ . Moreover for each pair of ordinals  $\gamma_1 \leq \overline{\gamma}_1$  and  $\gamma_2 \leq \overline{\gamma}_2$ , we define three arrows  $h_1(\gamma_1, \gamma_2) : Y_1 \rightarrow Z$ ,  $h_2(\gamma_1, \gamma_2) : Y_1 \rightarrow Z$ , and  $h_3(\gamma_1, \gamma_2) : Y_2 \rightarrow Z$  by:*

$$\begin{aligned} h_1(\gamma_1, \gamma_2) &= u_1(\gamma_1) \odot f_{11}^{(\gamma_2)}, & h_2(\gamma_1, \gamma_2) &= u_2(\overline{\gamma}_1) \odot f_{12}^{(\gamma_2)}, \\ & & \text{and } h_3(\gamma_1, \gamma_2) &= [u_1(\overline{\gamma}_1), u_2(\overline{\gamma}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle. \end{aligned}$$

*We claim that, if we let*

$$p := ((\overline{\gamma}_1, \overline{\gamma}_2), (h_1(\gamma_1, \gamma_2), h_2(\gamma_1, \gamma_2), h_3(\gamma_1, \gamma_2))_{\gamma_1 \leq \overline{\gamma}_1, \gamma_2 \leq \overline{\gamma}_2}), \quad (6.5)$$

$$\begin{array}{c}
\begin{array}{ccc}
FY \xrightarrow{\overline{F}[\langle\langle f_{11}^{(\gamma_2)}, f_{12}^{(\gamma_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} FX \xrightarrow{\overline{F}[u_1(\gamma_1), u_2(\gamma_1)]} FZ & & FY \xrightarrow{\overline{F}[\langle\langle f_{11}^{(\gamma_2)}, f_{12}^{(\gamma_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} FX \xrightarrow{\overline{F}[u_1(\overline{\gamma_1}), u_2(\overline{\gamma_1})]} FZ \\
\uparrow d_{11} \quad \sqsupseteq \quad \uparrow c_1 \quad \sqsupseteq \quad J\zeta \uparrow \cong & & \uparrow d_{12} \quad \sqsupseteq \quad \uparrow c_2 \quad \sqsupseteq \quad J\zeta \uparrow \cong \\
Y_1 \xrightarrow{f_{11}^{(\gamma_2)}} X_1 \xrightarrow{u_1(\gamma_1+1)} Z & & Y_1 \xrightarrow{f_{12}^{(\gamma_2+1)}} X_2 \xrightarrow{u_2(\overline{\gamma_1})} Z \\
\hline
& h_1(\gamma_1, \gamma_2) & h_2(\gamma_1, \gamma_2)
\end{array} \\
\\
\begin{array}{ccc}
FY \xrightarrow{\overline{F}[\langle\langle f_{11}^{(\overline{\gamma_2})}, f_{12}^{(\overline{\gamma_2})} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} FX \xrightarrow{\overline{F}[u_1(\overline{\gamma_1}), u_2(\overline{\gamma_1})]} FZ & & \\
\uparrow d_2 \quad \sqsupseteq \quad \uparrow c \quad \sqsupseteq \quad J\zeta \uparrow \cong & & \\
Y_2 \xrightarrow{\langle\langle f_{21}, f_{22} \rangle\rangle} X \xrightarrow{[u_1(\overline{\gamma_1}), u_2(\overline{\gamma_1})]} Z & & \\
\hline
& h_3(\gamma_1, \gamma_2) &
\end{array}
\end{array}$$

Figure 3: The progress measure  $p$  in (6.5) as diagrams.

then it is a progress measure for the following equational system.

$$\begin{aligned}
h_1 &=_{\mu} (J\zeta)^{-1} \odot \overline{F}[\text{id}_Z, \text{id}_Z] \odot \langle\langle h_1, h_2 \rangle\rangle, h_3 \odot d_{11} \in \mathcal{Kl}(T)(Y_1, Z) \\
h_2 &=_{\mu} (J\zeta)^{-1} \odot \overline{F}[\text{id}_Z, \text{id}_Z] \odot \langle\langle h_1, h_2 \rangle\rangle, h_3 \odot d_{12} \in \mathcal{Kl}(T)(Y_1, Z) \\
h_3 &=_{\nu} (J\zeta)^{-1} \odot \overline{F}[\text{id}_Z, \text{id}_Z] \odot \langle\langle h_1, h_2 \rangle\rangle, h_3 \odot d_2 \in \mathcal{Kl}(T)(Y_2, Z)
\end{aligned} \tag{6.6}$$

Note here that if  $h_1, h_2 \in \mathcal{Kl}(T)(Y_1, Z)$  satisfy the equations above, then their codomain join  $\langle\langle h_1, h_2 \rangle\rangle$  is always defined.

*Proof.* We check that  $p$  in (6.5) satisfies the axioms of progress measure (Definition 2.5). See also Figure 3.

- (1) **(Monotonicity)** We assume  $\gamma_1 \leq \gamma'_1$  and  $\gamma_2 \leq \gamma'_2$ . Then by Assumption 6.11.3 and that  $(f_{11}^{(\alpha)})_{\alpha \leq \overline{\alpha}}$  and  $(f_{12}^{(\alpha)})_{\alpha \leq \overline{\alpha}}$  are increasing sequence, we have:

$$h_1(\gamma_1, \gamma_2) = u_1(\gamma_1) \odot f_{11}^{(\gamma_2)} \sqsubseteq u_1(\gamma'_1) \odot f_{11}^{(\gamma'_2)} = h_1(\gamma'_1, \gamma'_2).$$

Hence monotonicity of  $h_1$  is proved. Monotonicity of  $h_2$  and  $h_3$  are proved similarly.

- (2) **( $\mu$ -variables, base case)** By Condition (2c) in Definition 6.12 and Condition (7) in Theorem 6.13, we have:

$$h_1(0, \gamma_2) = u_1(0) \odot f_{11}^{(\gamma_2)} = \perp \odot f_{11}^{(\gamma_2)} = \perp \quad \text{and} \quad h_2(\gamma_1, 0) = u_2(\overline{\gamma_1}) \odot f_{12}^{(0)} = u_2 \odot \perp = \perp.$$

- (3) **( $\mu$ -variables, step case)** Let  $\gamma_1 \leq \overline{\gamma_1}$  and  $\gamma_2 \leq \overline{\gamma_2}$ . We have the following (see also Figure 3).

$$\begin{aligned}
& h_1(\gamma_1 + 1, \gamma_2) \\
&= u_1(\gamma_1 + 1) \odot f_{11}^{(\gamma_2)} \\
&\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[u_1(\gamma_1), u_2(\gamma_1)] \odot c_1 \odot f_{11}^{(\gamma_2)} \\
&\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[u_1(\gamma_1), u_2(\overline{\gamma_1})] \odot c_1 \odot f_{11}^{(\gamma_2)} \\
&\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[u_1(\gamma_1), u_2(\overline{\gamma_1})] \odot \overline{F}[\langle\langle f_{11}^{(\gamma_2)}, f_{12}^{(\gamma_2)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11}
\end{aligned}$$

$$\begin{aligned}
 &= (J\zeta)^{-1} \odot \overline{F}[[u_1(\gamma_1), u_2(\overline{\gamma_1})] \odot \langle\langle f_{11}^{(\gamma_2)}, f_{12}^{(\gamma_2)} \rangle\rangle, [u_1(\gamma_1), u_2(\overline{\gamma_1})] \odot \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11} \\
 &= (J\zeta)^{-1} \odot \overline{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle u_1(\gamma_1) \odot f_{11}^{(\gamma_2)}, u_2(\overline{\gamma_1}) \odot f_{12}^{(\gamma_2)} \rangle\rangle, \\
 &\quad [u_1(\gamma_1), u_2(\overline{\gamma_1})] \odot \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11} \\
 &\sqsubseteq (J\zeta)^{-1} \odot \overline{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle u_1(\gamma_1) \odot f_{11}^{(\gamma_2)}, u_2(\overline{\gamma_1}) \odot f_{12}^{(\gamma_2)} \rangle\rangle, \\
 &\quad [u_1(\overline{\gamma_1}), u_2(\overline{\gamma_1})] \odot \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_{11} \\
 &= (J\zeta)^{-1} \odot \overline{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle h_1(\gamma_1, \gamma_2), h_2(\gamma_1, \gamma_2) \rangle\rangle, h_3(\gamma_1, \gamma_2)] \odot d_{11}.
 \end{aligned}$$

We can prove in a similar manner that there exists an ordinal  $\gamma'_1$  such that

$$h_2(\gamma_1, \gamma_2 + 1) \sqsubseteq \overline{F}[\langle\langle h_1(\gamma'_1, \gamma_2), h_2(\gamma_1, \gamma_2) \rangle\rangle, h_3(\gamma_1, \gamma_2)] \odot d_{12}.$$

- (4) ( **$\mu$ -variables, limit case**) Let  $\gamma_1$  be a limit ordinal such that  $\gamma_1 \leq \overline{\gamma_1}$ . By  $\overline{\gamma_1} = \overline{\beta_1} \leq \omega$  and Assumption 6.11.3, Kleisli composition in  $\mathcal{Kl}(T)$  is  $\gamma_1$ -continuous. Hence for each ordinal  $\gamma_2$ , we have:

$$\begin{aligned}
 h_1(\gamma_1, \gamma_2) &= u_1(\gamma_1) \odot f_{11}^{(\gamma_2)} \sqsubseteq \left( \bigsqcup_{\gamma'_1 < \gamma_1} u_1(\gamma'_1) \right) \odot f_{11}^{(\gamma_2)} \\
 &= \bigsqcup_{\gamma'_1 < \gamma_1} (u_1(\gamma'_1) \odot f_{11}^{(\gamma_2)}) = \bigsqcup_{\gamma'_1 < \gamma_1} h_1(\gamma'_1, \gamma_2).
 \end{aligned}$$

In a similar manner we can prove that for an ordinal  $\gamma_1$  and a limit ordinal  $\gamma_2$ ,

$$h_2(\gamma_1, \gamma_2) = \bigsqcup_{\gamma'_2 < \gamma_2} h_2(\gamma_1, \gamma'_2).$$

- (5) ( **$\nu$ -variables**) Similarly to the step cases of  $\mu$ -variables, we can prove that for ordinals  $\gamma_1 \leq \overline{\gamma_1}$  and  $\gamma_2 \leq \overline{\gamma_2}$  we have:

$$h_3(\gamma_1, \gamma_2) \sqsubseteq (J\zeta)^{-1} \odot \overline{F}[\langle\langle h_1(\overline{\gamma_1}, \overline{\gamma_2}), h_2(\overline{\gamma_1}, \overline{\gamma_2}) \rangle\rangle, h_3(\overline{\gamma_1}, \overline{\gamma_2})] \odot d_2.$$

Hence  $p$  is a progress measure for the equational system (6.6).  $\square$

**Sublemma 6.16.** *We assume Assumption 6.11 and the assumptions in Theorem 6.13. Let  $f : Y \rightarrow X$  be a forward,  $\bar{\alpha}$ -bounded simulation with dividing from  $\mathcal{X} = ((X_1, X_2), c, s)$  and  $\mathcal{Y} = ((Y_1, Y_2), d, t)$ . Let  $(v_1^{\text{sol}}, v_2^{\text{sol}}, v_3^{\text{sol}})$  be the solution of equational system (6.6) in Sublemma 6.15. Then we have:*

$$[\text{id}_Z, \text{id}_Z] \odot \langle\langle v_1^{\text{sol}}, v_2^{\text{sol}} \rangle\rangle \sqsubseteq \text{tr}^{\text{B}}(d_1) \quad \text{and} \quad v_3^{\text{sol}} \sqsubseteq \text{tr}^{\text{B}}(d_2) \quad (6.7)$$

where  $\text{tr}^{\text{B}}(d_1) : Y_1 \rightarrow Z$  and  $\text{tr}^{\text{B}}(d_2) : Y_2 \rightarrow Z$  are defined as  $\text{tr}^{\text{B}}(c_1)$  and  $\text{tr}^{\text{B}}(c_2)$ .

*Proof.* It is easy to see that the equational system in (6.6) is equivalent to the following equational system, in the sense that  $w_1^{\text{sol}} = (v_1^{\text{sol}}, v_2^{\text{sol}})$  and  $w_2^{\text{sol}} = v_3^{\text{sol}}$ .

$$\begin{aligned}
 w_1 &=_{\mu} \left( \begin{array}{l} (J\zeta)^{-1} \odot \overline{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}, w_{12} \rangle\rangle, w_2] \odot d_{11}, \\ (J\zeta)^{-1} \odot \overline{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}, w_{12} \rangle\rangle, w_2] \odot d_{12} \end{array} \right) \in (\mathcal{Kl}(T)(Y_1, Z))^2 \\
 w_2 &=_{\nu} (J\zeta)^{-1} \odot \overline{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}, w_{12} \rangle\rangle, w_2] \odot d_2 \in \mathcal{Kl}(T)(Y_2, Z)
 \end{aligned} \quad (6.8)$$

Here  $(\mathcal{Kl}(T)(Y_1, Z))^2$  is equipped with the product order, and  $w_{11}$  and  $w_{12}$  denote the first and the second component of  $w_1 \in (\mathcal{Kl}(T)(Y_1, Z))^2$  respectively.

By completeness of progress measure (Theorem 2.7.2), there exists a progress measure  $q = ((\alpha), (w_1(\alpha) = (w_{11}(\alpha), w_{12}(\alpha)), w_2(\alpha))_{\alpha \leq \bar{\alpha}})$  for (6.8) such that  $(w_{11}(\bar{\alpha}), w_{12}(\bar{\alpha})) = w_1^{\text{sol}} = (v_1^{\text{sol}}, v_2^{\text{sol}})$  and  $w_2(\bar{\alpha}) = w_2^{\text{sol}} = v_3^{\text{sol}}$ .

For each  $\alpha \leq \bar{\alpha}$ , we define  $v'_1(\alpha) : Y_1 \rightarrow Z$  and  $v'_2(\alpha) : Y_2 \rightarrow Z$  by  $v'_1(\alpha) = [\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\alpha), w_{12}(\alpha) \rangle\rangle$  and  $v'_2(\alpha) = w_2(\alpha)$ .

In what follows, we show that  $p' := ((\bar{\alpha}), (v'_1(\alpha), v'_2(\alpha))_{\alpha \leq \bar{\alpha}})$  is a progress measure for the equational system that defines  $\text{tr}^{\text{B}}(d_1) : Y_1 \rightarrow Z$  and  $\text{tr}^{\text{B}}(d_2) : Y_2 \rightarrow Z$  (see (5.2) in Theorem 5.6).

- (1) **(Monotonicity)** By the monotonicity of  $w_1(\alpha)$  and  $w_2(\alpha)$ ,  $v'_1(\alpha)$  and  $v'_2(\alpha)$  are also monotone.
- (2) **( $\mu$ -variables, base case)** We have  $(w_{11}(0), w_{12}(0)) = w_1(0) = (\perp, \perp)$  by the definition. Hence by Condition (7) of Theorem 6.13, we have:

$$v'_1(0) = [\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(0), w_{12}(0) \rangle\rangle = [\text{id}_Z, \text{id}_Z] \odot \langle\langle \perp, \perp \rangle\rangle = \perp.$$

- (3) **( $\mu$ -variables, step case)** For an ordinal  $\alpha \leq \bar{\alpha}$ , we have:

$$\begin{aligned} v'_1(\alpha + 1) &= [\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\alpha + 1), w_{12}(\alpha + 1) \rangle\rangle \\ &\sqsubseteq [\text{id}_Z, \text{id}_Z] \odot \langle\langle (J\zeta)^{-1} \odot \bar{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\alpha), w_{12}(\alpha) \rangle\rangle, w_2] \odot d_{11}, \\ &\quad (J\zeta)^{-1} \odot \bar{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\alpha), w_{12}(\alpha) \rangle\rangle, w_2] \odot d_{12} \rangle\rangle \\ &= (J\zeta)^{-1} \odot \bar{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\alpha), w_{12}(\alpha) \rangle\rangle, w_2] \odot [\text{id}_{FY}, \text{id}_{FY}] \odot \langle\langle d_{11}, d_{12} \rangle\rangle \\ &\sqsubseteq (J\zeta)^{-1} \odot \bar{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\alpha), w_{12}(\alpha) \rangle\rangle, w_2] \odot d_1 \\ &= (J\zeta)^{-1} \odot \bar{F}[v'_1(\alpha), v'_2(\alpha)] \odot d_1. \end{aligned}$$

- (4) **( $\mu$ -variables, limit case)** For a limit ordinal  $\alpha \leq \bar{\alpha}$ , we have:

$$\begin{aligned} v'_1(\alpha) &= [\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\alpha), w_{12}(\alpha) \rangle\rangle \\ &\sqsubseteq [\text{id}_Z, \text{id}_Z] \odot \langle\langle \bigsqcup_{\beta < \alpha} w_{11}(\beta), \bigsqcup_{\beta < \alpha} v'_{12}(\beta) \rangle\rangle \\ &= \bigsqcup_{\beta < \alpha} [\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\beta), w_{12}(\beta) \rangle\rangle \\ &= \bigsqcup_{\beta < \alpha} v'_1(\beta). \end{aligned}$$

- (5) **( $\nu$ -variables)** For an ordinal  $\alpha \leq \bar{\alpha}$ , there exists an ordinal  $\beta \leq \bar{\alpha}$  such that:

$$\begin{aligned} v'_2(\alpha) &= w_2(\alpha) \\ &\sqsubseteq (J\zeta)^{-1} \odot \bar{F}[[\text{id}_Z, \text{id}_Z] \odot \langle\langle w_{11}(\beta), w_{12}(\beta) \rangle\rangle, w_2(\beta)] \odot d_2 \\ &= (J\zeta)^{-1} \odot \bar{F}[v'_1(\beta), v'_2(\beta)] \odot d_2. \end{aligned}$$

Hence  $p' = ((\bar{\alpha}), (v'_1(\alpha), v'_2(\alpha))_{\alpha})$  is a progress measure and by soundness of progress measures (Theorem 2.7.1), we have (6.7).  $\square$

*Proof (Lemma 6.14).* Let  $E_{\mathcal{X}}$  be the equational system that defines  $\text{tr}^{\text{B}}(c_1)$  and  $\text{tr}^{\text{B}}(c_2)$  (see (5.2) in Theorem 5.6). Let  $p_{\mathcal{X}} = ((\bar{\beta}_1), (u_1(\beta_1), u_2(\beta_1))_{\beta_1 \leq \bar{\beta}_1})$  be the progress measure in Sublemma 6.15. By Sublemma 6.15 and soundness of progress measures (Theorem 2.7.1), we have:

$$\begin{aligned} u_1(\bar{\beta}_1) \odot f_{11}^{(\bar{\alpha}_1)} &\sqsubseteq v_1^{\text{sol}}, & u_2(\bar{\beta}_1) \odot f_{12}^{(\bar{\alpha}_1)} &\sqsubseteq v_2^{\text{sol}}, \\ & & \text{and} & [u_1(\bar{\beta}_1), u_2(\bar{\beta}_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle &\sqsubseteq v_3^{\text{sol}}. \end{aligned} \quad (6.9)$$

By Sublemma 6.16, we have:

$$[\text{id}_Z, \text{id}_Z] \odot [v_1^{\text{sol}}, v_2^{\text{sol}}] \sqsubseteq \text{tr}^{\text{B}}(d_1) \quad \text{and} \quad v_3^{\text{sol}} \sqsubseteq \text{tr}^{\text{B}}(d_1). \quad (6.10)$$

Therefore we have:

$$\begin{aligned} [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot \langle\langle f_{11}, f_{12} \rangle\rangle &= [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot \langle\langle f_{11}^{(\overline{\alpha_1})}, f_{12}^{(\overline{\alpha_1})} \rangle\rangle \quad (\text{by Definition 6.12}) \\ &= [u_1(\overline{\beta_1}), u_2(\overline{\beta_1})] \odot \langle\langle f_{11}^{(\overline{\alpha_1})}, f_{12}^{(\overline{\alpha_1})} \rangle\rangle \quad (\text{by definition}) \\ &= [\text{id}_Z, \text{id}_Z] \odot \langle\langle u_1(\overline{\beta_1}) \odot f_{11}^{(\overline{\alpha_1})}, u_2 \odot f_{12}^{(\overline{\alpha_1})} \rangle\rangle \\ &\sqsubseteq [\text{id}_Z, \text{id}_Z] \odot \langle\langle v_1^{\text{sol}}, v_2^{\text{sol}} \rangle\rangle \quad (\text{by (6.9)}) \\ &\sqsubseteq \text{tr}^{\text{B}}(d_1) \quad (\text{by (6.10)}). \end{aligned}$$

In a similar manner, we can prove:

$$[\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle \sqsubseteq \text{tr}^{\text{B}}(d_2).$$

These conclude the proof.  $\square$

*Proof (Theorem 6.13).* We have:

$$\begin{aligned} &\text{tr}^{\text{B}}(\mathcal{X}) \\ &= [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot s \quad (\text{by definition}) \\ &\sqsubseteq [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot [\langle\langle f_{11}, f_{12} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot t \quad (\text{by Condition (1) in Definition 6.12}) \\ &\sqsubseteq [[\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot \langle\langle f_{11}, f_{12} \rangle\rangle, [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle] \odot t \\ &\sqsubseteq [\text{tr}^{\text{B}}(d_1), \text{tr}^{\text{B}}(d_2)] \odot t \quad (\text{by Lemma 6.14}) \\ &= \text{tr}^{\text{B}}(\mathcal{Y}). \quad (\text{by definition}) \end{aligned}$$

This concludes the proof.  $\square$

We have thus obtained a sound simulation notion. The proposition below shows that soundness theorem (Theorem 6.13) applies to the combinations of monads and functors in Definition 5.1–5.2.

**Proposition 6.17.** *The combinations of  $\mathcal{P}$  and  $F_\Sigma$ , and  $\mathcal{G}$  and  $F_{\mathbf{A}}$ , respectively, satisfy the assumptions in Assumption 6.11 and Theorem 6.13.  $\square$*

Therefore by Theorem 6.13 and Theorem 5.6, by regarding NBTAs and PBWAs as Büchi  $(T, F)$ -systems as in Example 5.5, we can obtain fair simulation notions for these systems whose soundness comes for free.

A problem here is that the coalgebraic definition in Definition 6.12 requires a dividing  $d_{11}, d_{12}: Y_1 \rightarrow \overline{F}Y$  of  $d_1: Y_1 \rightarrow \overline{F}Y$ . Intuitively this is to divide the simulator's “resources” of transitions into two parts, one for the challenger's *non-accepting* states and the other for *accepting* states.

To describe an intuition, we hereby interpret the notion of dividing for NBTAs and PBWAs, with respect to the correspondence in Example 5.5. For the NBTA  $\mathcal{Y}$  in Def. 3.11, a dividing is understood as a pair  $\delta_{\mathcal{Y},11}, \delta_{\mathcal{Y},12}: Y_1 \rightarrow \mathcal{P}(\coprod_{\sigma \in \Sigma} X^{|\sigma|})$  of functions such that  $\delta_{\mathcal{Y},11}(x) \cup \delta_{\mathcal{Y},12}(y) = \delta_{\mathcal{Y}}(y)$  for each  $y \in Y_1$ . If  $\mathcal{Y}$  is the PBWA in Def. 4.12, then a dividing is understood as a pair  $M_{\mathcal{Y},11}(a), M_{\mathcal{Y},12}(a) \in [0, 1]^{Y_1 \times Y}$  of matrices such that  $M_{\mathcal{Y},11}(a) + M_{\mathcal{Y},12}(a) = M_{\mathcal{Y},1}(a)$  for each  $a \in \mathbf{A}$ .

This dividing requirement is naturally inherited by the resulting concrete simulation notions for NBTAs and PBWAs. Unfortunately finding such “resource allocation” is a challenge in practice; additionally, insistence on such allocation being *static* is overly restrictive, as we will later see in Example 6.19.

The following definition is more desirable in this respect; it indeed yields Definition 3.11 and 4.12—the concrete simulation notions that we have introduced earlier—as its instances. Note that the following definition is *not sound* in the general sense of Theorem 6.13 (see Example 6.27 for a counterexample). The rest of the paper is devoted to finding special cases in which it is sound.

**Definition 6.18** (fair simulation without dividing). In the setting of Definition 6.12, a (*forward,  $\bar{\alpha}$ -bounded*) *fair simulation without dividing* is defined almost the same way as one with dividing in Definition 6.12, except that Condition (2) is replaced by the following.

(2') The components  $f_{11}: Y_1 \rightarrow X_1$  and  $f_{12}: Y_1 \rightarrow X_2$  come with *approximation sequences*

$$f_{11}^{(0)} \sqsubseteq f_{11}^{(1)} \sqsubseteq \dots \sqsubseteq f_{11}^{(\bar{\alpha})} : Y_1 \rightarrow X_1, \quad \text{and} \quad f_{12}^{(0)} \sqsubseteq f_{12}^{(1)} \sqsubseteq \dots \sqsubseteq f_{12}^{(\bar{\alpha})} : Y_1 \rightarrow X_2,$$

that satisfies 2a, 2c and 2e in Definition 6.12 and the following two conditions.

(b') ( $f_{11}^{(\alpha)}$ ) For each ordinal  $\alpha$  such that  $\alpha \leq \bar{\alpha}$ , the inequality (6.11) below holds.

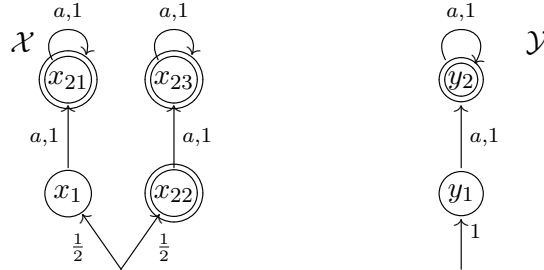
$$\begin{array}{ccc} FY & \xrightarrow{\bar{F}[\langle\langle f_{11}^{(\alpha)}, f_{12}^{(\alpha)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\ d_1 \uparrow & \sqsupseteq & \uparrow c_1 \\ Y_1 & \xrightarrow{f_{11}^{(\alpha)}} & X_1 \end{array} \quad (6.11)$$

(d') ( $f_{12}^{(\alpha)}$ , the step case) For each  $\alpha < \bar{\alpha}$ , the inequality (6.12) below holds.

$$\begin{array}{ccc} FY & \xrightarrow{\bar{F}[\langle\langle f_{11}^{(\alpha)}, f_{12}^{(\alpha)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle]} & FX \\ d_1 \uparrow & \sqsupseteq & \uparrow c_2 \\ Y_1 & \xrightarrow{f_{12}^{(\alpha+1)}} & X_2 \end{array} \quad (6.12)$$

The following example shows why this definition is more desirable.

**Example 6.19.** Let  $\mathcal{X} = ((X_1, X_2), c, s)$  and  $\mathcal{Y} = ((Y_1, Y_2), d, t)$  be Büchi  $(\mathcal{G}, \{a\} \times (\_))$ -systems that model PBWAs illustrated below.



It is easy to see that they exhibit language inclusion. We define  $f : Y \rightarrow X$  by

$$f(y_1)(\{x\}) = \begin{cases} \frac{1}{2} & (x \in \{x_1, x_{22}\}) \\ 0 & (\text{otherwise}) \end{cases} \quad \text{and} \quad f(y_2)(\{x\}) = \begin{cases} \frac{1}{2} & (x \in \{x_{21}, x_{23}\}) \\ 0 & (\text{otherwise}). \end{cases}$$



Then  $f$  is a fair simulation *without* dividing from  $\mathcal{X}$  to  $\mathcal{Y}$ . In contrast,  $f$  is not a fair simulation *with* dividing. In fact, there exists no fair simulation with dividing from  $\mathcal{X}$  to  $\mathcal{Y}$ .

In what follows we seek for conditions under which this desirable fair simulation notion (without dividing, Definition 6.18) turns out to be sound. In Section 6.3 we study the nondeterministic setting, and in Section 6.4 we study the probabilistic setting. The identified conditions and soundness arguments are rather different between Section 6.3 and Section 6.4.

**6.3. Circumventing Dividing: the Nondeterministic Case.** For  $T = \mathcal{P}$  we show that a simulation *without* dividing yields one *with* dividing. Therefore in the nondeterministic setting fair simulations without dividing are sound unconditionally. Here we exploit the *idempotency* property of  $T = \mathcal{P}$ —*one can copy resources as many times as one likes*.

**Proposition 6.20** (soundness under idempotency). *Under Assumption 6.11, let us assume that each arrow  $f : X \rightarrow Y$  in  $\mathcal{Kl}(T)$  is idempotent, that is, the codomain join  $\langle\langle f, f \rangle\rangle : X \rightarrow Y + Y$  necessarily exists and we have  $[\text{id}_Y, \text{id}_Y] \odot \langle\langle f, f \rangle\rangle = f$ .*

- (1) *A simulation without dividing yields one with dividing, with the dividing  $d_{11} = d_{12} = d_1$ .*
- (2) *Under the assumptions of Theorem 6.13, a simulation without dividing witnesses trace inclusion.* □

*Proof.* The item (1) is immediate from the definitions of simulation with dividing and one without dividing (Definition 6.12 and Definition 6.18). The item (2) follows from (1) and soundness of forward fair simulation with dividing (Theorem 6.13). □

**Lemma 6.21.** *Arrows in  $\mathcal{Kl}(\mathcal{P})$  are idempotent. Hence all the conditions in Proposition 6.20 are satisfied by  $T = \mathcal{P}$  and  $F = F_\Sigma$ , where  $\Sigma$  is a ranked alphabet.* □

There is still a gap between the simulation notion in Definition 6.18 (defined by inequalities) and that in Definition 3.11 (defined by an equational system). The gap is filled by another specific property of  $\mathcal{Kl}(\mathcal{P})$ —*reversibility*. It is much like in the following “must” predicate transformers.

**Lemma 6.22.** *Let  $f : B \rightarrow C$  be an arrow in  $\mathcal{Kl}(\mathcal{P})$ . We define  $\square_f : \mathcal{Kl}(\mathcal{P})(A, C) \rightarrow \mathcal{Kl}(\mathcal{P})(A, B)$  by*

$$\square_f(g)(x) := \{y \in B \mid f(y) \subseteq g(x)\}.$$

*Then we have the following:*

- (1)  $f \odot \square_f(g) \sqsubseteq g$
- (2)  $\forall h : A \rightarrow B. f \odot h \sqsubseteq g \Rightarrow h \sqsubseteq \square_f(g)$ . □

The construction  $\square_f$  is used to essentially “reverse” the arrows  $c_1$  and  $c_2$  on the right of the diagrams (6.2–6.3). This allows to separate variables  $f_{11}$ ,  $f_{12}$ ,  $f_{21}$  and  $f_{22}$  alone and yield a (proper) equational system as in (6.13) below.

**Proposition 6.23.** *Let  $g_1^{sol} : Y_1 \rightarrow X_1$ ,  $g_2^{sol} : Y_1 \rightarrow X_2$ ,  $g_3^{sol} : Y_2 \rightarrow X_1$  and  $g_4^{sol} : Y_2 \rightarrow X_2$  be the solution of the following equational system.*

$$\begin{aligned} g_1 &=_{\nu} \square_{c_1}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_1) \in \mathcal{Kl}(\mathcal{P})(Y_1, X_1) \\ g_2 &=_{\mu} \square_{c_2}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_1) \in \mathcal{Kl}(\mathcal{P})(Y_1, X_2) \\ g_3 &=_{\nu} \square_{c_1}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_2) \in \mathcal{Kl}(\mathcal{P})(Y_2, X_1) \\ g_4 &=_{\nu} \square_{c_2}(\overline{F}[\langle\langle g_1, g_2 \rangle\rangle, \langle\langle g_3, g_4 \rangle\rangle] \odot d_2) \in \mathcal{Kl}(\mathcal{P})(Y_2, X_2) \end{aligned} \tag{6.13}$$

Let  $g^{sol} = [\langle\langle g_1^{sol}, g_2^{sol} \rangle\rangle, \langle\langle g_3^{sol}, g_4^{sol} \rangle\rangle] : Y \rightarrow X$ . Then  $s \sqsubseteq g^{sol} \odot t$  if and only if there is a fair  $\bar{\alpha}$ -bounded simulation without dividing (Definition 6.18) from  $\mathcal{X}$  to  $\mathcal{Y}$  for some ordinal  $\bar{\alpha}$ .

*Proof.* As in Definition 6.12, we write  $f_{ji} : Y_j \rightarrow X_i$  for the domain and codomain restriction of  $f : Y \rightarrow X$ .

( $\Rightarrow$ ). Assume  $s \sqsubseteq g^{sol} \odot t$ . By completeness of progress measure (Theorem 2.7.2), there exists a progress measure  $g = ((\bar{\alpha}), (g_i(\alpha))_{1 \leq i \leq 4, \alpha \leq \bar{\alpha}})$  such that  $g_i^{sol} = g_i(\bar{\alpha})$  for each  $i$ . We define two sequences  $(f_{11}^{(\alpha)} : Y_1 \rightarrow X_1)_{\alpha \leq \bar{\alpha}}$  and  $(f_{12}^{(\alpha)} : Y_1 \rightarrow X_2)_{\alpha \leq \bar{\alpha}}$  by  $f_{11}^{(\alpha)} = g_1(\alpha)$  and  $f_{12}^{(\alpha)} = g_2(\alpha)$ . We define  $f : Y \rightarrow X$  by  $f = g^{sol}$ . We show that  $f$  is a fair simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  whose approximation sequences are given by  $(f_{11}^{(\alpha)})_{\alpha \leq \bar{\alpha}}$  and  $(f_{12}^{(\alpha)})_{\alpha \leq \bar{\alpha}}$ .

We first show that  $f$  satisfies Condition (1) in Definition 6.12. We have:

$$\begin{aligned}
c \odot f &= c \odot [\langle\langle f_{11}, f_{12} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \\
&\sqsubseteq c \odot [\langle\langle \square_{c_1}(\bar{F}f \odot d_1), \square_{c_2}(\bar{F}f \odot d_1) \rangle\rangle, \langle\langle \square_{c_1}(\bar{F}f \odot d_2), \square_{c_2}(\bar{F}f \odot d_2) \rangle\rangle] \\
&= [\langle\langle c_1 \odot \square_{c_1}(\bar{F}f \odot d_1), c_2 \odot \square_{c_2}(\bar{F}f \odot d_1) \rangle\rangle, \langle\langle c_1 \odot \square_{c_1}(\bar{F}f \odot d_2), c_2 \odot \square_{c_2}(\bar{F}f \odot d_2) \rangle\rangle] \\
&\sqsubseteq [\langle\langle \bar{F}f \odot d_1, \bar{F}f \odot d_1 \rangle\rangle, \langle\langle \bar{F}f \odot d_2, \bar{F}f \odot d_2 \rangle\rangle] \quad (\text{by Lemma 6.22.1}) \\
&= [\bar{F}f \odot d_1, \bar{F}f \odot d_2] \quad (\text{by Lemma 6.21}) \\
&= \bar{F}f \odot d. \tag{6.14}
\end{aligned}$$

Moreover, by the assumption, we have  $s \sqsubseteq f \odot t$ .

Next we show that  $f$  satisfies the Condition (2') in Definition 6.18. Note that  $g = ((\bar{\alpha}), (g_i(\alpha))_{1 \leq i \leq 4, \alpha \leq \bar{\alpha}})$  satisfies the axioms of progress measure (Definition 2.5). It is immediate that this implies that conditions 2a, 2c and 2e in Definition 6.12 are satisfied. Moreover in a similar manner to (6.14) above, this implies that 2'b' and 2'd' in Definition 6.18 are also satisfied. Therefore  $f$  is a fair  $\bar{\alpha}$ -bounded simulation without dividing.

( $\Leftarrow$ ). Conversely, let  $f : Y \rightarrow X$  be a fair simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  whose approximation sequences are given by  $(f_{11}^{(\alpha)})_{\alpha \leq \bar{\alpha}}$  and  $(f_{12}^{(\alpha)})_{\alpha \leq \bar{\alpha}}$ . For each  $\alpha \leq \bar{\alpha}$ , we define arrows  $g_1(\alpha) : Y_1 \rightarrow X_1$ ,  $g_2(\alpha) : Y_1 \rightarrow X_2$ ,  $g_3(\alpha) : Y_2 \rightarrow X_1$  and  $g_4(\alpha) : Y_2 \rightarrow X_2$ , by  $g_1(\alpha) = f_{11}^{(\alpha)}$ ,  $g_2(\alpha) = f_{12}^{(\alpha)}$ ,  $g_3(\alpha) = f_{21}$  and  $g_4(\alpha) = f_{22}$ . Then by using Lemma 6.22.2, we can easily show that Condition (1) in Definition 6.12 and Condition (2') in Definition 6.18, together with monotonicity of  $f_{11}^{(\alpha)}$  and  $f_{12}^{(\alpha)}$ , imply that  $g$  satisfies the axioms of a progress measure (Definition 2.5) with respect to the equational system (6.13).  $\square$

By Proposition 6.23 and that  $\bar{F}$  and  $\odot$  in  $\mathcal{Kl}(\mathcal{P})$  are  $\alpha$ -continuous for an arbitrary limit ordinal  $\alpha$ , it is not hard to translate Proposition 6.20 into Theorem 3.12.

*Proof (Theorem 3.12).* Let  $\mathcal{X}$  and  $\mathcal{Y}$  be NBTAs and

$$\mathcal{X} = ((X_1, X_2), c, s) \text{ and } \mathcal{Y} = ((Y_1, Y_2), d, t)$$

be the corresponding Büchi  $(\mathcal{P}, F_\Sigma)$ -systems (see Example 5.5).

Note that for each  $A, B \in \mathbf{Sets}$ , a function  $\Delta_{A,B} : \mathcal{P}(A \times B) \rightarrow \mathcal{Kl}(\mathcal{P})(B, A)$  that is defined by  $\Delta_{A,B}(S)(b) := \{a \in A \mid (a, b) \in S\}$  is a bijection. It is easy to see that for the functions in Definition 3.11 and Lemma 6.22, we have the following (recall that

$$F_\Sigma A = \coprod_{\sigma \in \Sigma} A^{|\sigma|}.$$

$$\begin{aligned} \Delta_{X_i, Y}(\square_{\mathcal{X}, i}(S)) &= \square_{c_i}(\Delta_{FX, B'}(S)) && \text{for } i \in 1, 2 \text{ and } S \subseteq \coprod_{\sigma \in \Sigma} X^{|\sigma|} \times Y \\ \Delta_{FX, Y_j}(\diamond_{\mathcal{Y}, j}(T)) &= (\Delta_{FX, FY}(T)) \odot d_j && \text{for } j \in \{1, 2\} \text{ and } T \subseteq \coprod_{\sigma \in \Sigma} X^{|\sigma|} \times \coprod_{\sigma \in \Sigma} Y^{|\sigma|} \\ \Delta_{FX, FY}(\wedge_\Sigma(U)) &= \overline{F}(\Delta_{X, Y}(U)) && \text{for } U \subseteq X \times Y \end{aligned}$$

It is also easy to see that  $R \subseteq Y \times X$  satisfies  $\forall x \in I_{\mathcal{X}}. \exists y \in I_{\mathcal{Y}}. (x, y) \in R$  if and only if  $s \sqsubseteq \Delta_{Y, X}(R) \odot t$ .

Hence together with Proposition 6.20 and Proposition 6.23, we have:

$R \subseteq X \times Y$  is a fair simulation from  $\mathcal{X}$  to  $\mathcal{Y}$  in the sense of Definition 3.11

$$\Leftrightarrow \forall x \in I_{\mathcal{X}}. \exists y \in I_{\mathcal{Y}}. (x, y) \in R,$$

$$\text{and } R \subseteq u_1^{\text{sol}} \cup u_2^{\text{sol}} \cup u_3^{\text{sol}} \cup u_4^{\text{sol}} \text{ where } u_1^{\text{sol}}, \dots, u_4^{\text{sol}} \text{ are the solution of (3.1)}$$

$$\Leftrightarrow s \sqsubseteq \Delta_{Y, X}(R) \odot t,$$

$$\text{and } \Delta_{X, Y}(R) \sqsubseteq [\langle\langle g_1^{\text{sol}}, g_2^{\text{sol}} \rangle\rangle, \langle\langle g_3^{\text{sol}}, g_4^{\text{sol}} \rangle\rangle] \text{ where } g_1^{\text{sol}}, \dots, g_4^{\text{sol}} \text{ are the solution of (6.13)}$$

$$\Leftrightarrow \text{there is a fair } \bar{\alpha}\text{-bounded simulation without dividing from } \mathcal{X} \text{ to } \mathcal{Y} \text{ for some } \bar{\alpha}$$

$$\Rightarrow \text{language inclusion, that is, } L(\mathcal{X}) \subseteq L(\mathcal{Y}).$$

This concludes the proof.  $\square$

The results here in Section 6.3 are axiomatic—with idempotency and reversibility—and they apply to monads other than  $\mathcal{P}$ . One example is the *lift monad*  $\mathcal{L} = 1 + (\_)$ , which is used for potential nontermination (see [HJS07] for example).

**6.4. Circumventing Dividing: the Probabilistic Case.** We turn to the probabilistic setting and prove Theorem 4.13. The strategy for  $T = \mathcal{P}$  does not work here, because  $\mathcal{G}$  lacks idempotency (see Proposition 6.20). We shall rely on other restrictions: from trees to words; and finite state spaces on the simulating side.

We start with an axiomatic development.

**Proposition 6.24.** *Besides Assumption 6.11 and the assumptions in Theorem 6.13, assume  $d_1 = \overline{F}(\text{id}_{Y_1} + \perp_{Y_2, Y_2}) \odot d_1$ . Then existence of a fair simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  implies trace inclusion.*

*Proof.* Let  $(f_{ij} : Y_i \rightarrow X_j)_{i, j \in \{1, 2\}}$  be a fair simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$ . Let  $f_{11}^{(0)} \sqsubseteq f_{11}^{(1)} \sqsubseteq \dots \sqsubseteq f_{11}^{(\bar{\alpha})} = f_{11}$  and  $f_{12}^{(0)} \sqsubseteq f_{12}^{(1)} \sqsubseteq \dots \sqsubseteq f_{12}^{(\bar{\alpha})} = f_{12}$  be the approximation sequences.

Recall that  $\text{tr}^B(c_1) : X_1 \rightarrow Z$  and  $\text{tr}^B(c_2) : X_2 \rightarrow Z$  are given by the solutions  $u_1^{\text{sol}}$  and  $u_2^{\text{sol}}$  of the following equational system. (Theorem 5.6).

$$\begin{aligned} u_1 &=_{\mu} (J\zeta)^{-1} \odot \overline{F}[u_1, u_2] \odot c_1 \in \mathcal{Kl}(T)(X_1, Z) \\ u_2 &=_{\nu} (J\zeta)^{-1} \odot \overline{F}[u_1, u_2] \odot c_2 \in \mathcal{Kl}(T)(X_2, Z) \end{aligned} \tag{6.15}$$

By completeness of progress measure (Theorem 2.7.2), there exists a progress measure

$$p_{\mathcal{X}} = ((\bar{\beta}), (u_1(\beta) : X_1 \rightarrow Z, u_2(\beta) : X_2 \rightarrow Z)_{\beta \leq \bar{\beta}})$$

for (6.15) such that  $u_1(\bar{\beta}) = \text{tr}^B(c_1)$  and  $u_2(\bar{\beta}) = \text{tr}^B(c_2)$ .

We define two ordinals  $\overline{\gamma}_1$  and  $\overline{\gamma}_2$  by  $\overline{\gamma}_1 = \overline{\beta}$  and  $\overline{\gamma}_2 = \overline{\alpha}$ . For each  $\gamma_1 \leq \overline{\gamma}_1$  and  $\gamma_2 \leq \overline{\gamma}_2$ , we define  $h_1(\gamma_1, \gamma_2) : Y_1 \rightarrow Z$  and  $h_2(\gamma_1, \gamma_2) : Y_1 \rightarrow Z$

$$h_1(\gamma_1, \gamma_2) = u_1(\gamma_1) \odot f_{11}^{(\gamma_2)}, \quad \text{and} \quad h_2(\gamma_1, \gamma_2) = u_2(\gamma_1) \odot f_{12}^{(\gamma_2)},$$

We shall prove  $h_1(\gamma_1, \gamma_2) = h_2(\gamma_1, \gamma_2) = \perp$  by transfinite induction on  $\gamma_1$  and  $\gamma_2$ .

**(base case).** If  $\gamma_1 = 0$ , for each  $\gamma_2 \leq \overline{\gamma}_2$  we have:

$$\begin{aligned} h_1(\gamma_1, \gamma_2) &= u_1(0) \odot f_{11}^{(\gamma_2)} && \text{(by definition)} \\ &= \perp \odot f_{11}^{(\gamma_2)} && (p_{\mathcal{X}} \text{ is a progress measure}) \\ &= \perp && \text{(by Condition (7) in Theorem 6.13)} \end{aligned}$$

Similarly, if  $\gamma_2 = 0$ , for each  $\gamma_1 \leq \overline{\gamma}_1$  we have:  $h_2(\gamma_1, \gamma_2) = \perp$ .

**(step case).** Assume we have  $h_1(\gamma_1, \gamma_2 + 1) = h_2(\gamma_1, \gamma_2 + 1) = \perp$ . Then we have:

$$\begin{aligned} &h_1(\gamma_1 + 1, \gamma_2 + 1) \\ &= u_1(\gamma_1 + 1) \odot f_{11}^{(\gamma_2 + 1)} && \text{(by definition)} \\ &\sqsubseteq J\zeta^{-1} \odot \overline{F}[u_1(\gamma_1), u_2(\gamma_1)] \odot c_1 \odot f_{11}^{(\gamma_2 + 1)} && (p_{\mathcal{X}} \text{ is a progress measure}) \\ &\sqsubseteq J\zeta^{-1} \odot \overline{F}[u_1(\gamma_1), u_2(\gamma_1)] \odot \overline{F}[\langle\langle f_{11}^{(\gamma_2 + 1)}, f_{12}^{(\gamma_2 + 1)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_1 \\ & && (f \text{ is a forward fair simulation}) \\ &= J\zeta^{-1} \odot \overline{F}[u_1(\gamma_1), u_2(\gamma_1)] \odot \overline{F}[\langle\langle f_{11}^{(\gamma_2 + 1)}, f_{12}^{(\gamma_2 + 1)} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot \overline{F}(\text{id} + \perp) \odot d_1 \\ & && \text{(by the assumption)} \\ &= J\zeta^{-1} \odot \overline{F}\left[[u_1(\gamma_1), u_2(\gamma_1)] \odot \langle\langle f_{11}^{(\gamma_2 + 1)}, f_{12}^{(\gamma_2 + 1)} \rangle\rangle\right] \odot \text{id}, \\ & && [u_1(\gamma_1), u_2(\gamma_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle \odot \perp \Big] \odot d_1 \\ &= J\zeta^{-1} \odot \overline{F}\left[[\text{id}, \text{id}] \odot \langle\langle u_1(\gamma_1) \odot f_{11}^{(\gamma_2 + 1)}, u_2(\gamma_1) \odot f_{12}^{(\gamma_2 + 1)} \rangle\rangle, \perp\right] \odot d_1 \\ &= J\zeta^{-1} \odot \overline{F}\left[[\text{id}, \text{id}] \odot \langle\langle h_1(\gamma_1, \gamma_2 + 1), h_2(\gamma_1, \gamma_2 + 1) \rangle\rangle, \perp\right] \odot d_1 && \text{(by definition)} \\ &= J\zeta^{-1} \odot \overline{F}\left[[\text{id}, \text{id}] \odot \langle\langle \perp, \perp \rangle\rangle, \perp\right] \odot d_1 && \text{(by the induction hypothesis)} \\ &= \perp. \end{aligned}$$

Similarly, if  $h_1(\gamma_1 + 1, \gamma_2) = h_2(\gamma_1 + 1, \gamma_2) = \perp$ , then we have:  $h_2(\gamma_1 + 1, \gamma_2 + 1) = \perp$ .

**(limit case).** Assume that  $\gamma_1$  is a limit ordinal and we have  $h_1(\gamma'_1, \gamma_2) = h_2(\gamma'_1, \gamma_2) = \perp$  for each  $\gamma'_1 < \gamma_1$ , and  $h_1(\gamma_1, \gamma'_2) = h_2(\gamma_1, \gamma'_2) = \perp$  for each  $\gamma'_2 < \gamma_2$ .

We first prove  $h_1(\gamma_1, \gamma_2) = \perp$ .

$$\begin{aligned} h_1(\gamma_1, \gamma_2) &= u_1(\gamma_1) \odot f_{11}^{(\gamma_2)} && \text{(by definition)} \\ &\sqsubseteq \left( \bigsqcup_{\gamma'_1 < \gamma_1} u_1(\gamma'_1) \right) \odot f_{11}^{(\gamma_2)} && (p_{\mathcal{X}} \text{ is a progress measure}) \\ &= \bigsqcup_{\gamma'_1 < \gamma_1} \left( u_1(\gamma'_1) \odot f_{11}^{(\gamma_2)} \right) && (\mathcal{Kl}(T) \text{ is } \mathbf{Cppo}\text{-enriched}) \end{aligned}$$

$$\begin{aligned}
 &= \bigsqcup_{\gamma'_1 < \gamma_1} h_1(\gamma'_1, \gamma_2) && \text{(by definition)} \\
 &= \perp && \text{(by IH)}
 \end{aligned}$$

We next prove  $h_2(\gamma_1, \gamma_2) = \perp$ . If  $\gamma_2$  is zero or a successor ordinal, we can prove  $h_2(\gamma_1, \gamma_2) = \perp$  much like the base case and the step cases in the above. If  $\gamma_2$  is a limit ordinal, then in a similar manner to the above, we have  $h_2(\gamma_1, \gamma_2) = \perp$ . Hence we have  $h_1(\gamma_1, \gamma_2) = h_2(\gamma_1, \gamma_2) = \perp$ .

Similarly, for a limit ordinal  $\gamma_2 \leq \overline{\gamma_2}$  such that  $h_1(\gamma'_1, \gamma_2) = h_2(\gamma'_1, \gamma_2) = \perp$  for each  $\gamma'_1 < \gamma_1$ , and  $h_1(\gamma_1, \gamma'_2) = h_2(\gamma_1, \gamma'_2) = \perp$  for each  $\gamma'_2 < \gamma_2$ , we have  $h_1(\gamma_1, \gamma_2) = h_2(\gamma_1, \gamma_2) = \perp$ .

Hence we have  $h_1(\gamma_1, \gamma_2) = h_2(\gamma_1, \gamma_2) = \perp$  for each  $\gamma_1 \leq \overline{\gamma_1}$  and  $\gamma_2 \leq \overline{\gamma_2}$ . Therefore we have:

$$[\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot \langle\langle f_{11}, f_{12} \rangle\rangle = \perp \sqsubseteq \text{tr}^{\text{B}}(d_1). \quad (6.16)$$

We define  $h_3 : Y_2 \rightarrow Z$  by  $h_3 = [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle$ . Then we have:

$$\begin{aligned}
 h_3 &= [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_1)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle && \text{(by definition)} \\
 &\sqsubseteq J\zeta^{-1} \odot \overline{F} \left[ [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_1)] \odot [f_{11}, f_{12}], [\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_1)] \odot [f_{21}, f_{22}] \right] \odot d_2 \\
 & && \text{(similarly to the above)} \\
 &= J\zeta^{-1} \odot \overline{F}[\perp, h_3] \odot d_2 && \text{(by definition and discussions above)} \\
 &\sqsubseteq J\zeta^{-1} \odot \overline{F}[l_1^{(1)}(h_3), h_3] \odot d_2.
 \end{aligned}$$

Here  $l_1^{(1)} : Y_1 \rightarrow Z$  denotes the first interim solution in Definition 2.2. Note that  $\text{tr}^{\text{B}}(d_2) : Y_2 \rightarrow Z$  is the greatest fixed point of the following function.

$$g \mapsto J\zeta^{-1} \odot \overline{F}[l_1^{(1)}(g), g] \odot d_2$$

Hence by the Knaster-Tarski theorem, we have

$$[\text{tr}^{\text{B}}(c_1), \text{tr}^{\text{B}}(c_2)] \odot \langle\langle f_{21}, f_{22} \rangle\rangle = h_3 \sqsubseteq \text{tr}^{\text{B}}(d_2). \quad (6.17)$$

By (6.16) and (6.17), in a similar manner to the proof of Theorem 6.13, we can prove  $\text{tr}^{\text{B}}(\mathcal{X}) \sqsubseteq \text{tr}^{\text{B}}(\mathcal{Y})$ .  $\square$

The following (non-coalgebraic) lemma states that if  $T = \mathcal{G}$ ,  $F = \mathbf{A} \times (\_)$  and the state space of  $\mathcal{Y}$  is finite, then we can modify  $\mathcal{Y}$  so that the assumption in Proposition 6.24 holds without changing its language. The modification is derived from the well-known *fairness* result on Markov chains (see [BK08, Chapter 10] for example). Concretely, this result states that a nonaccepting state  $\bigcirc$  from which an accepting state is reachable in a positive probability can be changed into an accepting state  $\odot$ . The proof uses the notion of *bottom strongly connected component*.

**Lemma 6.25.** *Let  $\mathbf{A}$  be a countable set and  $\mathcal{Y} = ((Y_1, Y_2), d, t)$  be a Büchi  $(\mathcal{G}, \mathbf{A} \times (\_))$ -system such that  $Y_1$  and  $Y_2$  are finite sets. Let  $y_{>0} \in Y_1$  be a state such that  $d(y_{>0})(\mathbf{A} \times Y_2) > 0$ . We define a Büchi  $(\mathcal{G}, \mathbf{A} \times (\_))$ -system  $\mathcal{Y}' = ((Y'_1, Y'_2), d', t')$  by:  $Y'_1 = Y_1 \setminus \{y_{>0}\}$ ,  $Y'_2 = Y_2 + \{y_{>0}\}$ ,  $d' = d$  and  $t' = t$ . Note that  $d'$  and  $t'$  are well-defined because  $Y'_1 + Y'_2 = Y_1 + Y_2$ .*

*Then we have  $[\text{tr}^{\text{B}}(d_1), \text{tr}^{\text{B}}(d_2)] = [\text{tr}^{\text{B}}(d'_1), \text{tr}^{\text{B}}(d'_2)]$ , and moreover,  $\text{tr}^{\text{B}}(\mathcal{Y}) = \text{tr}^{\text{B}}(\mathcal{Y}')$ .*

*Proof.* Let  $Y = Y_1 + Y_2$ . Note that  $Y$  is a finite set equipped with a discrete  $\sigma$ -algebra.

The Büchi  $(\mathcal{G}, \mathbf{A} \times (\_))$ -system  $\mathcal{Y} = ((Y_1, Y_2), t, d)$  induces a Markov chain  $\mathcal{M}_{\mathcal{Y}}$  such that the state space is defined by  $Y_{\perp} = Y + \{\perp\}$  and the transition function  $\tau_{\mathcal{Y}} : Y_{\perp} \times Y_{\perp} \rightarrow [0, 1]$  is given by

$$\tau_{\mathcal{Y}}(y, y') = \begin{cases} \sum_{a \in \mathbf{A}} d(y)(\{(a, y')\}) & (y, y' \in Y) \\ 1 - \sum_{y' \in Y} \sum_{a \in \mathbf{A}} d(y)(\{(a, y')\}) & (y \in Y, y' = \perp) \\ 1 & (y = y' = \perp) \\ 0 & (\text{otherwise}). \end{cases}$$

A Markov chain  $\mathcal{M}_{\mathcal{Y}'}$  is defined similarly.

A subset  $B \subseteq Y$  is called a *strongly connected component* (SCC for short) if for all  $y, y' \in S$ , there exist  $y_0, y_1, \dots, y_n$  such that  $y_0 = y$ ,  $y_n = y'$  and  $\tau_{\mathcal{Y}}(y_i, y_{i+1}) > 0$  for each  $i$ . An SCC  $B$  is called a *bottom strongly connected component* (BSCC for short) if  $\tau_{\mathcal{Y}}(y, y') = 0$  for each  $y \in B$  and  $y' \notin B$ . For more details, see [BK08] for example.

For  $Y' \subseteq Y$ , we write  $\Pr(y \models \text{GF}Y')$  for the probability in which a state in  $Y'$  is visited infinitely often on  $\mathcal{M}_{\mathcal{Y}}$  from  $y \in Y$ . By Theorem 5.6.2b, we have:

$$\begin{aligned} [\text{tr}^{\text{B}}(d_1), \text{tr}^{\text{B}}(d_2)](y)(\mathbf{A}^{\omega}) &= \Pr(y \models \text{GF}Y_2), \quad \text{and} \\ [\text{tr}^{\text{B}}(d'_1), \text{tr}^{\text{B}}(d'_2)](y)(\mathbf{A}^{\omega}) &= \Pr(y \models \text{GF}(Y_2 + \{y_{>0}\})). \end{aligned}$$

We define  $U, U' \subseteq Y$  by

$$\begin{aligned} U &:= \bigcup \{B \subseteq Y \mid B \text{ is a BSCC and } B \cap Y_2 \neq \emptyset\}, \quad \text{and} \\ U' &:= \bigcup \{B \subseteq Y \mid B \text{ is a BSCC and } B \cap (Y_2 + \{y_{>0}\}) \neq \emptyset\}. \end{aligned}$$

We write  $\Pr(y \models \text{F}U)$  for a probability in which a state in  $U$  is reached in  $\mathcal{M}_{\mathcal{Y}}$ . It is known that  $\Pr(y \models \text{GF}Y_2) = \Pr(y \models \text{F}U)$  (see [BK08, Corollary 10.34] for example). Similarly, we have  $\Pr(y \models \text{GF}(Y_2 + \{y_{>0}\})) = \Pr(y \models \text{F}U')$ .

Assume that  $y_{>0} \in B$  for some BSCC  $B$  in  $\mathcal{M}_{\mathcal{Y}}$ . As  $B$  is a BSCC, it has no outgoing transition, on one hand. On the other hand, by  $d(y_{>0})(\mathbf{A} \times Y_2) > 0$ ,  $y_{>0}$  has an accepting successor state. Hence by the definition of  $U$ , we have  $B \cap Y_2 \neq \emptyset$  and this implies that  $U = U'$ .

If  $y_{>0} \notin B$  for any BSCC  $B$ , then by the definitions of  $U$  and  $U'$  we have  $U = U'$ .

Therefore in both cases, for each  $y \in Y$ , we have:

$$\begin{aligned} [\text{tr}^{\text{B}}(d_1), \text{tr}^{\text{B}}(d_2)](y)(\mathbf{A}^{\omega}) &= \Pr(y \models \text{GF}Y_2) \\ &= \Pr(y \models \text{F}U) \\ &= \Pr(y \models \text{F}U') \\ &= \Pr(y \models \text{GF}(Y_2 + \{y_{>0}\})) \\ &= [\text{tr}^{\text{B}}(d'_1), \text{tr}^{\text{B}}(d'_2)](y)(\mathbf{A}^{\omega}). \end{aligned}$$

It remains to prove  $[\text{tr}^{\text{B}}(d_1), \text{tr}^{\text{B}}(d_2)](y)(A) = [\text{tr}^{\text{B}}(d'_1), \text{tr}^{\text{B}}(d'_2)](y)(A)$  for each measurable set  $A \subseteq \mathbf{A}^{\omega}$ . To this end, by Carathéodory's extension theorem (see [ADD00] for example), it suffices to prove  $[\text{tr}^{\text{B}}(d_1), \text{tr}^{\text{B}}(d_2)](y)(w\mathbf{A}^{\omega}) = [\text{tr}^{\text{B}}(d'_1), \text{tr}^{\text{B}}(d'_2)](y)(w\mathbf{A}^{\omega})$  for each  $w \in \mathbf{A}^*$ .

We inductively define a function  $\chi_{\mathcal{Y}} : Y \times \mathbf{A}^* \rightarrow \mathcal{G}Y$  by

$$\begin{aligned} \chi_{\mathcal{Y}}(y, \varepsilon)(\{y'\}) &= \begin{cases} 1 & (y = y') \\ 0 & (\text{otherwise}) \end{cases} \quad \text{and} \\ \chi_{\mathcal{Y}}(y, aw)(\{y'\}) &= \sum_{y'' \in Y} d(y)(\{(a, y'')\}) \cdot \chi_{\mathcal{Y}}(y'', w)(\{y'\}) \end{aligned}$$

where  $a \in \mathbf{A}$  and  $w \in \mathbf{A}^*$ .

Then for each  $y \in Y$  and  $w \in \mathbf{A}^*$ , we have:

$$\begin{aligned} [\text{tr}^{\mathbf{B}}(d_1), \text{tr}^{\mathbf{B}}(d_2)](y)(w\mathbf{A}^\omega) &= \sum_{y' \in Y} \chi_{\mathcal{Y}}(y, w)(y') \cdot [\text{tr}^{\mathbf{B}}(d_1), \text{tr}^{\mathbf{B}}(d_2)](y')(\mathbf{A}^\omega) \\ &= \sum_{y' \in Y} \chi_{\mathcal{Y}}(y, w)(y') \cdot [\text{tr}^{\mathbf{B}}(d'_1), \text{tr}^{\mathbf{B}}(d'_2)](y')(\mathbf{A}^\omega) \\ &= [\text{tr}^{\mathbf{B}}(d'_1), \text{tr}^{\mathbf{B}}(d'_2)](y)(w\mathbf{A}^\omega). \end{aligned}$$

By Carathéodory's extension theorem, this implies  $[\text{tr}^{\mathbf{B}}(d_1), \text{tr}^{\mathbf{B}}(d_2)](y)(A) = [\text{tr}^{\mathbf{B}}(d'_1), \text{tr}^{\mathbf{B}}(d'_2)](y)(A)$  for each measurable set  $A$ . Hence we have  $\text{tr}^{\mathbf{B}}(\mathcal{Y}) = \text{tr}^{\mathbf{B}}(\mathcal{Y}')$ .  $\square$

With Lemma 6.25 discharging its assumptions, Proposition 6.24 easily yields Theorem 4.13 as follows.

*Proof (Theorem 4.13).* Let  $\mathcal{X} = ((X_1, X_2), c, s)$  and  $\mathcal{Y} = ((Y_1, Y_2), d, t)$ . By using the bijective correspondence between probabilistic matrices and arrows in  $\mathcal{Kl}(\mathcal{G})(X, Y)$  where  $X$  and  $Y$  are equipped with discrete  $\sigma$ -algebras, we can easily see that a forward fair matrix simulation  $A \in [0, 1]^{Y \times X}$  from  $\mathcal{X}$  to  $\mathcal{Y}$  (Definition 4.12) exists if and only if a fair simulation  $f : Y \rightarrow X$  without dividing from  $\mathcal{X}$  to  $\mathcal{Y}$  (Definition 6.18) exists (see also [UH14, UH17]).

We define  $Y_{12} \subseteq Y_1$  by

$$Y_{12} = \{y \in Y_1 \mid \exists y_0, \dots, y_n \in Y, y = y_0, y_n \in Y_2, \forall i. d(y_i)(\mathbf{A} \times \{y_{i+1}\}) > 0\}.$$

As  $Y_1$  is finite,  $Y_{12}$  is also finite. We define a Büchi  $(\mathcal{G}, \mathbf{A} \times (\_))$ -system  $\mathcal{Y}' = ((Y'_1, Y'_2), d', t')$  by  $Y'_1 = Y_1 \setminus Y_{12}$ ,  $Y'_2 = Y_2 + Y_{12}$ ,  $d' = d$  and  $t' = t$ . As  $Y_{12}$  is finite, by repeatedly applying Lemma 6.25, we have  $\text{tr}^{\mathbf{B}}(\mathcal{Y}') = \text{tr}^{\mathbf{B}}(\mathcal{Y})$ .

It is easy to see that  $f$  is also a fair simulation without dividing from  $\mathcal{X}$  to  $\mathcal{Y}'$ . Moreover, by its definition,  $\mathcal{Y}'$  satisfies  $\text{tr}^{\mathbf{B}}(d'_1) = \perp$ . Therefore by Proposition 6.24, we have  $\text{tr}^{\mathbf{B}}(\mathcal{X}) \sqsubseteq \text{tr}^{\mathbf{B}}(\mathcal{Y}')$ . Hence  $\text{tr}^{\mathbf{B}}(\mathcal{X}) \sqsubseteq \text{tr}^{\mathbf{B}}(\mathcal{Y})$  holds.  $\square$

It is still open whether the restriction from trees to words is necessary. We note that an analogous statement to Lemma 6.25 does not hold for Büchi  $(\mathcal{G}, \mathbf{A} \times (\_) \times (\_))$ -systems, which can be regarded as probabilistic Büchi tree automata. A counterexample is as follows.

**Example 6.26.** We define a Büchi  $(\mathcal{G}, \{a\} \times (\_) \times (\_))$ -system  $\mathcal{Y} = ((Y_1, Y_2), d, t)$  by:

$$\begin{aligned} Y_1 &= (\{y_1\}, \mathcal{P}\{y_1\}), \quad Y_2 = (\{y_2\}, \mathcal{P}\{y_2\}), \\ d(y_1)(\{(a, y, y')\}) &= \begin{cases} \frac{1}{2} & (y = y' = y_1) \\ \frac{1}{4} & (y = y_1, y' = y_2 \text{ or } y = y' = y_2) \\ 0 & (\text{otherwise}), \end{cases} \\ d(y_2)(\{(a, y, y')\}) &= \begin{cases} 1 & (y = y' = y_2) \\ 0 & (\text{otherwise}), \end{cases} \quad \text{and} \quad t(*) (\{y\}) = \begin{cases} 1 & (y = y_1) \\ 0 & (\text{otherwise}). \end{cases} \end{aligned}$$

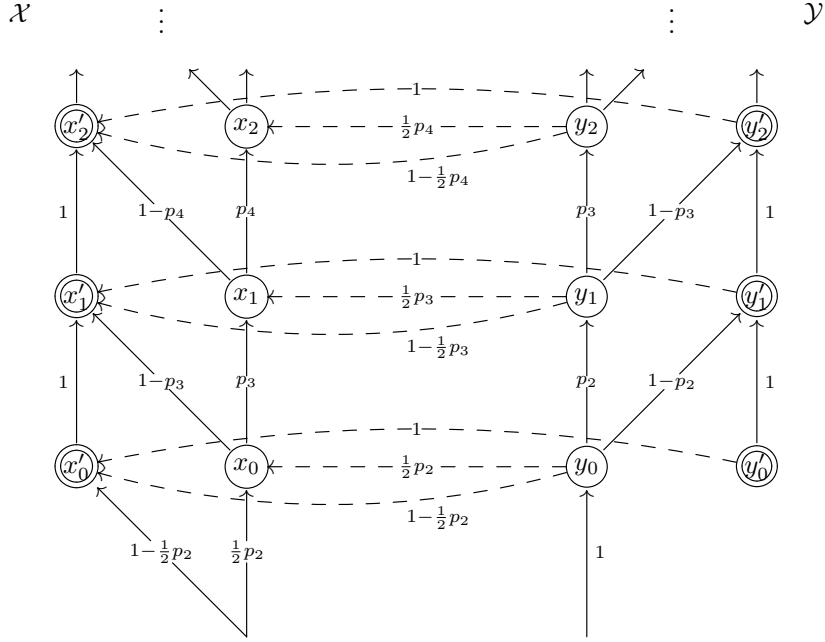


Figure 4: Büchi  $(\mathcal{G}, \{a\} \times (\_))$ -systems  $\mathcal{X}$  and  $\mathcal{Y}$  (whose transitions are denoted by solid lines), and a forward fair simulation (denoted by dashed lines). Labels are omitted.

Note that  $d(y_1)(\{a\} \times Y_2 \times Y_2) = \frac{1}{2} > 0$ . Note also that the carrier set of the final  $(\{a\} \times (\_) \times (\_))$ -coalgebra is given by a singleton  $\{*\}$ . It is not hard to see that  $\text{tr}^B(d_1)(y_1)(\{*\}) = \frac{1}{2}$ . In contrast, if we define a Büchi  $(\mathcal{G}, \{a\} \times (\_) \times (\_))$ -system  $\mathcal{Y}' = ((Y'_1, Y'_2), d', t')$  by  $Y'_1 = \emptyset$ ,  $Y'_2 = \{y_1, y_2\}$ ,  $d' = d$  and  $t = t'$ , then we have  $\text{tr}^B(d'_1)(y_1)(\{*\}) = 1$ .

In contrast, it turns out that the finiteness restriction of  $\mathcal{Y}$  in Theorem 4.13 is strict: an example below shows that without it soundness fails.

**Example 6.27.** Let  $\mathcal{X} = ((X_1, X_2), c, s)$  and  $\mathcal{Y} = ((Y_1, Y_2), d, t)$  be Büchi  $(\mathcal{G}, \{a\} \times (\_))$ -systems that are illustrated as PBWAs in Figure 4. For each  $i \in \omega$ , we define  $p_i \in [0, 1]$  by  $p_i = 1 - 1/i^2$ . We define a family  $f = (f_{ij} : Y_i \rightarrow X_j)_{i,j \in \{1,2\}}$  of Kleisli arrows as follows (see also dashed lines in Figure 4):

- $f_{11}(y_i)(\{x_j\}) = \frac{1}{2}p_{i+2}$  if  $j = i$  and 0 otherwise;
- $f_{12}(y_i)(\{x'_j\}) = 1 - \frac{1}{2}p_{i+2}$  if  $j = i$  and 0 otherwise;
- $f_{21}(y'_i)(\{x_j\}) = 0$ ; and
- $f_{22}(y'_i)(\{x'_j\}) = 1$  if  $j = i$  and 0 otherwise.

Moreover, for an ordinal  $\alpha \in \{0, 1\}$ , we define Kleisli arrows  $f_{11}(\alpha) : Y_1 \rightarrow Y_1$  and  $f_{12}(\alpha) : Y_1 \rightarrow X_2$  by  $f_{11}(0) = \perp$ ,  $f_{11}(1) = f_{11}$ ,  $f_{12}(0) = \perp$  and  $f_{12}(1) = f_{12}$ .



Then the following inequalities hold.

$$\begin{aligned}
 s &\sqsubseteq [\langle\langle f_{11}, f_{12} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot t \\
 c \odot \langle\langle f_{11}(0), f_{12}(1) \rangle\rangle &\sqsubseteq \overline{F} [\langle\langle f_{11}(0), f_{12}(0) \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_1 \\
 c \odot \langle\langle f_{11}(1), f_{12}(1) \rangle\rangle &\sqsubseteq \overline{F} [\langle\langle f_{11}(1), f_{12}(1) \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_1 \\
 c \odot \langle\langle f_{21}, f_{22} \rangle\rangle &\sqsubseteq \overline{F} [\langle\langle f_{11}, f_{12} \rangle\rangle, \langle\langle f_{21}, f_{22} \rangle\rangle] \odot d_2
 \end{aligned}$$

Therefore  $f$  is a forward fair simulation without dividing (Definition 6.18) from  $\mathcal{X}$  to  $\mathcal{Y}$ .

In contrast, we also have:

$$\begin{aligned}
 \text{tr}(\mathcal{X})(\{a^\omega\}) &= \frac{1}{2}p_2 \cdot (1 - \prod_{i \in \omega} p_{i+3}) + (1 - \frac{1}{2}p_2) \cdot 1 = 1 - \frac{1}{2} \prod_{i \in \omega} p_{i+2} \quad \text{and} \\
 \text{tr}(\mathcal{Y})(\{a^\omega\}) &= 1 - \prod_{i \in \omega} p_{i+2}.
 \end{aligned}$$

As  $\prod_{i \in \omega} p_{i+2} = \frac{1}{2} > 0$ , we have  $\text{tr}(\mathcal{X}) \not\leq \text{tr}(\mathcal{Y})$ . Therefore language inclusion fails.

## 7. CONCLUSIONS AND FUTURE WORK

We defined notions of fair simulation for two types of transition systems with Büchi acceptance conditions, namely: nondeterministic Büchi tree automata (NBTAs) and probabilistic Büchi word automata (PBWAs, with an additional finiteness assumption on the simulating side).

The simulation notion for NBTAs is defined in terms of fixed-point equations (Definition 3.11). The resulting notion is almost the same as the one in [vB08], except that infinite state spaces are allowed in our definition because of the fixed-point formulation. In contrast, the fair simulation notion for PBWAs is new to the best of our knowledge. It is a combination of a notion of matrix simulation [UH17] and a notion of (lattice-theoretic) progress measure [Jur00, HSC16].

These simulation notions originate from categorical backgrounds that are built upon a categorical characterization of parity languages developed in [USH16]. Using the theory in [USH16], we have introduced a categorical simulation notion called (forward) fair simulation with dividing (Definition 6.12) and proved its soundness.

Our soundness proof for the categorical notion of fair simulation with dividing is mathematically clean, and we can easily obtain sound simulation notions for NBTAs and PBWAs by specializing the categorical notion. However, the resulting notions inherit dividing requirement, and it is a big disadvantage. For NBTAs and PBWAs, we have shown that by using properties that are specific to these systems, and by imposing a finite-state restriction to the simulating side for PBWAs, the dividing requirement can be lifted.

**7.1. Future Work.** Generalization from the Büchi condition to the *parity* one is certainly what we aim at next. It is already not very clear how our coalgebraic definition with dividing (Section 6.2) would generalize: for example, in case of parity automata, there is little sense in comparing the priority of the challenger’s state with that of the simulator’s. It is even less clear how to circumvent dividing.

Aside from fair simulation, notions of *delayed simulation* are known for Büchi automata [EWS05, FW06]: they are subject to slightly different “fairness” constraints. Accommodating them in the current setting is another future work.

On decidability and complexity, while the obtained simulation notion for NBTAs is decidable if the state spaces are finite, the decidability of that for PBWAs is still open even for finite-state systems. Obviously it is one of possible directions of future work.

We are also interested in automatic discovery of simulations—via mathematical programming for example—and its implementation. In this direction of future work we will be based on our previous work [UH14,UH17]. Another direction is to use the current results for *program verification*—where the `Integer` type makes problems inherently infinitary—exploiting our non-combinatorial presentation by equational systems that allows infinite state spaces. We could do so automatically by synthesizing a symbolic simulation or, interactively on a proof assistant.

**Acknowledgments.** Thanks are due to Shunsuke Shimizu, Kenta Cho, Eugenia Sironi, and the anonymous referees, for discussions and comments. The authors are supported by ERATO HASUO Metamathematics for Systems Design Project (No. JPMJER1603), JST, and Grant-in-Aid No. 15KT0012, JSPS. Natsuki Urabe is supported by Grant-in-Aid No. 16J08157 for JSPS Fellows.

#### REFERENCES

- [ABH<sup>+</sup>12] Jirí Adámek, Filippo Bonchi, Mathias Hülsbusch, Barbara König, Stefan Milius, and Alexandra Silva. A coalgebraic perspective on minimization and determinization. In Lars Birkedal, editor, *FoSSaCS*, volume 7213 of *Lect. Notes Comp. Sci.*, pages 58–73. Springer, 2012.
- [Adá74] Jirí Adámek. Free algebras and automata realizations in the language of categories. *Comment. Math. Univ. Carolin.*, 15:589602, 1974.
- [ADD00] Robert B. Ash and Catherine Doleans-Dade. *Probability and measure theory Second Edition*. Academic Press, 2000.
- [AK79] Jirí Adámek and Václav Koubek. Least fixed point of a functor. *J. Comput. Syst. Sci.*, 19(2):163–178, 1979.
- [AMV11] Jirí Adámek, Stefan Milius, and Jiri Velebil. Elgot theories: a new perspective on the equational properties of iteration. *Mathematical Structures in Computer Science*, 21(2):417–480, 2011.
- [AN01] André Arnold and Damian Niwiński. *Rudiments of  $\mu$ -Calculus*. Studies in Logic and the Foundations of Mathematics. Elsevier, Amsterdam, 2001.
- [BC03] Vincent D. Blondel and Vincent Canterini. Undecidable problems for probabilistic automata of fixed dimension. *Theory Comput. Syst.*, 36(3):231–245, 2003.
- [BK08] Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.
- [BMSZ14] Filippo Bonchi, Stefan Milius, Alexandra Silva, and Fabio Zanasi. How to kill epsilons with a dagger - A coalgebraic take on systems with algebraic label structure. In Marcello M. Bonsangue, editor, *Coalgebraic Methods in Computer Science - 12th IFIP WG 1.3 International Workshop, CMCS 2014, Colocated with ETAPS 2014, Grenoble, France, April 5-6, 2014, Revised Selected Papers*, volume 8446 of *Lecture Notes in Computer Science*, pages 53–74. Springer, 2014.
- [Bor94] Francis Borceux. *Handbook of Categorical Algebra.*, volume 2. Cambridge University Press, Cambridge, 11 1994.
- [CC79] Patrick Cousot and Radhia Cousot. Constructive versions of Tarski’s fixed point theorems. *Pacific Journal of Mathematics*, 82(1):43–57, 1979.
- [CHS14] Arnaud Carayol, Axel Haddad, and Olivier Serre. Randomization in automata on infinite trees. *ACM Trans. Comput. Log.*, 15(3):24:1–24:33, 2014.
- [Cîr10] Corina Cîrstea. Generic infinite traces and path-based coalgebraic temporal logics. *Electr. Notes Theor. Comput. Sci.*, 264(2):83–103, 2010.
- [Cîr13] Corina Cîrstea. From branching to linear time, coalgebraically. In David Baelde and Arnaud Carayol, editors, *Proceedings Workshop on Fixed Points in Computer Science, FICS 2013, Turino, Italy, September 1st, 2013.*, volume 126 of *EPTCS*, pages 11–27, 2013.

- [CKS92] Rance Cleaveland, Marion Klein, and Bernhard Steffen. Faster model checking for the modal mu-calculus. In Gregor von Bochmann and David K. Probst, editors, *Computer Aided Verification, Fourth International Workshop, CAV '92, Montreal, Canada, June 29 - July 1, 1992, Proceedings*, volume 663 of *Lecture Notes in Computer Science*, pages 410–422. Springer, 1992.
- [Doo94] J.L. Doob. *Measure Theory*. Graduate Texts in Mathematics. Springer New York, 1994.
- [EWS05] Kousha Etessami, Thomas Wilke, and Rebecca A. Schuller. Fair simulation relations, parity games, and state space reduction for Büchi automata. *SIAM J. Comput.*, 34(5):1159–1175, 2005.
- [FW06] Carsten Fritz and Thomas Wilke. Simulation relations for alternating parity automata and parity games. In Oscar H. Ibarra and Zhe Dang, editors, *Developments in Language Theory, 10th International Conference, DLT 2006, Santa Barbara, CA, USA, June 26-29, 2006, Proceedings*, volume 4036 of *Lecture Notes in Computer Science*, pages 59–70. Springer, 2006.
- [Gir82] Michele Giry. A categorical approach to probability theory. In *Proc. Categorical Aspects of Topology and Analysis*, volume 915 of *Lect. Notes Math.*, pages 68–85, 1982.
- [Has06] Ichiro Hasuo. Generic forward and backward simulations. In Christel Baier and Holger Hermanns, editors, *CONCUR*, volume 4137 of *Lecture Notes in Computer Science*, pages 406–420. Springer, 2006.
- [Her06] Horst Herrlich. *Axiom of Choice*. Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2006.
- [HJS07] Ichiro Hasuo, Bart Jacobs, and Ana Sokolova. Generic trace semantics via coinduction. *Logical Methods in Computer Science*, 3(4), 2007.
- [HKR02] Thomas A. Henzinger, Orna Kupferman, and Sriram K. Rajamani. Fair simulation. *Inf. Comput.*, 173(1):64–81, 2002.
- [HSC15] Ichiro Hasuo, Shunsuke Shimizu, and Corina Cîrstea. Lattice-theoretic progress measures and coalgebraic model checking (with appendices). *CoRR*, abs/1511.00346, 2015.
- [HSC16] Ichiro Hasuo, Shunsuke Shimizu, and Corina Cîrstea. Lattice-theoretic progress measures and coalgebraic model checking. In Rastislav Bodik and Rupak Majumdar, editors, *Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2016, St. Petersburg, FL, USA, January 20 - 22, 2016*, pages 718–732. ACM, 2016.
- [Jac04] Bart Jacobs. Trace semantics for coalgebras. *Electr. Notes Theor. Comput. Sci.*, 106:167–184, 2004.
- [Jac10] Bart Jacobs. From coalgebraic to monoidal traces. In *Coalgebraic Methods in Computer Science (CMCS 2010)*, volume 264 of *Elect. Notes in Theor. Comp. Sci.*, pages 125–140. Elsevier, Amsterdam, 2010.
- [Jac16] Bart Jacobs. *Introduction to Coalgebra: Towards Mathematics of States and Observation*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2016.
- [JL91] Bengt Jonsson and Kim Guldstrand Larsen. Specification and refinement of probabilistic processes. In *LICS*, pages 266–277. IEEE Computer Society, 1991.
- [JP06] Sudeep Juvekar and Nir Piterman. Minimizing generalized Büchi automata. In Thomas Ball and Robert B. Jones, editors, *Computer Aided Verification: 18th International Conference, CAV 2006, Seattle, WA, USA, August 17-20, 2006. Proceedings*, pages 45–58, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.
- [JSS15] Bart Jacobs, Alexandra Silva, and Ana Sokolova. Trace semantics via determinization. *J. Comput. Syst. Sci.*, 81(5):859–879, 2015.
- [Jur00] Marcin Jurdzinski. Small progress measures for solving parity games. In Horst Reichel and Sophie Tison, editors, *STACS*, volume 1770 of *Lecture Notes in Computer Science*, pages 290–301. Springer, 2000.
- [KK13] Henning Kerstan and Barbara König. Coalgebraic trace semantics for continuous probabilistic transition systems. *Logical Methods in Computer Science*, 9(4), 2013.
- [LV95] Nancy A. Lynch and Frits W. Vaandrager. Forward and backward simulations. I. Untimed systems. *Inf. Comput.*, 121(2):214–233, 1995.
- [LV96] Nancy A. Lynch and Frits W. Vaandrager. Forward and backward simulations. II. Timing based systems. *Inf. Comput.*, 128(1):1–25, 1996.
- [Mac98] S. Mac Lane. *Categories for the Working Mathematician*. Springer, Berlin, 2nd edition, 1998.
- [PT97] J. Power and H. Thielecke. Environments, continuation semantics and indexed categories. In M. Abadi and T. Ito, editors, *Theoretical Aspects of Computer Software*, number 1281 in *Lect. Notes Comp. Sci.*, pages 391–414. Springer, Berlin, 1997.

- [Rut00] J. J. M. M. Rutten. Universal coalgebra: a theory of systems. *Theor. Comp. Sci.*, 249:3–80, 2000.
- [TW<sup>+</sup>02] Wolfgang Thomas, Thomas Wilke, et al. *Automata, logics, and infinite games: a guide to current research*, volume 2500. Springer Science & Business Media, 2002.
- [UH14] Natsuki Urabe and Ichiro Hasuo. Generic forward and backward simulations III: quantitative simulations by matrices. In Paolo Baldan and Daniele Gorla, editors, *CONCUR 2014 - Concurrency Theory - 25th International Conference, CONCUR 2014, Rome, Italy, September 2-5, 2014. Proceedings*, volume 8704 of *Lecture Notes in Computer Science*, pages 451–466. Springer, 2014.
- [UH15] Natsuki Urabe and Ichiro Hasuo. Coalgebraic infinite traces and Kleisli simulations. In *Algebra and Coalgebra in Computer Science - 6th International Conference, CALCO 2015, Nijmegen, Netherlands, June 24-26, 2015. Proceedings*, 2015.
- [UH17] Natsuki Urabe and Ichiro Hasuo. Quantitative simulations by matrices. *Inf. Comput.*, 252:110–137, 2017.
- [UHH17] Natsuki Urabe, Masaki Hara, and Ichiro Hasuo. Categorical liveness checking by corecursive algebras. In *Proc. LICS 2017*, 2017. To appear.
- [USH16] Natsuki Urabe, Shunsuke Shimizu, and Ichiro Hasuo. Coalgebraic trace semantics for bueschi and parity automata. In Josée Desharnais and Radha Jagadeesan, editors, *27th International Conference on Concurrency Theory, CONCUR 2016, August 23-26, 2016, Québec City, Canada*, volume 59 of *LIPICs*, pages 24:1–24:15. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016.
- [Var95] Moshe Y. Vardi. An automata-theoretic approach to linear temporal logic. In Faron Moller and Graham M. Birtwistle, editors, *Banff Higher Order Workshop*, volume 1043 of *Lecture Notes in Computer Science*, pages 238–266. Springer, 1995.
- [vB08] Thomas von Bomhard. Minimization of tree automata. BSc thesis, Universität des Saarlandes, September 2008.
- [vBW05] Franck van Breugel and James Worrell. A behavioural pseudometric for probabilistic transition systems. *Theor. Comput. Sci.*, 331(1):115–142, 2005.
- [vG01] R. J. van Glabbeek. The linear time–branching time spectrum I; the semantics of concrete, sequential processes. In J. A. Bergstra, A. Ponse, and S. A. Smolka, editors, *Handbook of Process Algebra*, chapter 1, pages 3–99. Elsevier, 2001.